An evaluation of the predictive value of temperament tests carried out on Assistance Hearing Dogs.

A thesis submitted in partial fulfilment of the requirements for the

---

MSc by Research in Biology (animal behaviour)

at

University Of Lincoln

By

Hannah Plant

University Of Lincoln

2016

Abstract of thesis submitted in partial fulfilment of the requirements for the MSc by research in Biology

(Animal Behaviour)

Evaluation of the predictive value of temperament tests carried out on Assistance Hearing Dogs.

By

Hannah Plant

Assistance Hearing Dogs provide independence to people with a hearing impairment by alerting to household and danger sounds, and provide companionship in a range of public places. Hearing Dogs experience thorough socialising and training in order to meet high Assistance Dog standards. Present procedures in use by Hearing Dogs For Deaf People, UK involve an 8 months behavioural test, with the aim to predict each dogs' future training performance. Knowing the predictive validity of this test, with respect to task performance is of great importance. The sensitivity, specificity, and predictive validity of 8 month behaviour tests already in use by the UK charity, Hearing Dogs for Deaf People, for performance at 18 months were evaluated on 62 dogs over a two year period. Correlations between the two tests and significant differences in performance were assessed using Spearman's Rho and Wilcoxon matched pairs tests. The following 13 behavioural factors were analysed: social behaviour with adults, children and dogs, environmental behaviour, recovery rate, adaptability, vocal reactivity, motivation, trainability, frustration, chase, hunt and distraction. Convergent validity of both the 8 and 18 month behaviour tests was further examined where possible using a CBARQ questionnaire completed at 8 and 18 months by the socialisers and trainers of 12 of the dogs. Positive correlations were observed between all behavioural factors with the exception of distractibility and recovery rate, however overall the behavioural tests demonstrated a high sensitivity and a low specificity, signifying a poor predictive validity with respect to task performance. While these results partly reflect the small number of failures in each test, the findings may also be symptomatic of the continued socialising and training that all dogs, irrelevant of pass rates, undergo between the 8 and 18 month behaviour tests. The value of the 8 month test and its individual elements is critically appraised.

Acknowledgements

This research was partially supported by Hearing Dogs for Deaf People. I sincerely thank them for all their help and in particular Jenny Palser, for all her patience, support and encouragement throughout this two year thesis.

I would also like to express my heartfelt gratitude for the University of Lincoln and in particular Daniel Mills for giving me the opportunity to conduct this research and providing me with so much valuable advice and guidance throughout.

To James Serpell, I would also like to send my gratitude for his consent to use the CBARQ questionnaire which provided a large body of this work, and for his advice when scoring it.

To all my friends and family, you never let me give up. Thank you.

Contents

Figures and Tables

Figures

Page

Tables

List of acronyms

HD	Hearing Dogs For deaf people
CBARQ	Canine behavioural assessment and research Questionnaire
T1	8 month old Hearing Dog Behavioural Test
T2	18 month old Hearing Dog Behavioural test
CBARQ-1	Canine behavioural assessment and research Questionnaire given out to 8 month old dogs
CBARQ-2	Canine behavioural assessment and research Questionnaire given out to 8 month old dogs

Chapter 1

Hearing Dogs For Deaf People

Contents **Page**

1.1 Introduction- Hearing Dogs For Deaf People

Assistance dogs are extremely important for the recipients for whom they work. They provide a sense of security, companionship and have an important role in providing greater interactions between family members and the wider community (Hart et al 1996, Lane et al 1998, Guest and McNicholas 2006), as well as having an impact on a person's psychological and physical wellbeing (Sachs-Ericsson 2002, Rintala et al, 2002, Guest and Mcnicholas 2006, Lane et al 1998, Audrestch et al 2015). Assistance dogs help by fulfilling a variety of roles such as guiding the blind, assisting with physical disabilities for the disabled, providing a warning prior to an oncoming seizure, and alerting to other medical issues such as low blood sugar, or other psychiatric disabilities. Assistance Hearing Dogs alert their deaf handler to important household sounds and danger signals in work places and shops (Hart et al 1996, Guest and McNicholas 2006).

Hearing Dogs For Deaf people is an assistance dog charity (Registered charity number: England and Wales 293358, Scotland SCO40486). The charity was launched in 1982, inspired by existing hearing dog organisations already functioning in various parts of the United States. Since its beginning, to date, it has created approximately 1,600 working partnerships between deaf recipients and hearing dogs (Hearing dogs for deaf people, 2015). The dogs provide assistance to a deaf person by alerting them to sounds in the house, such as the alarm clock, the doorbell, the cooker timer, plus many more. The dogs alert by either touching the person with their nose or, if they are small, with their paw. After the dog has touched, the deaf handler asks 'what is it' with an accompanying hand signal. The dog then leads them to the relevant sound or lays down to indicate a danger sound such a smoke alarm or fire siren.

However just as importantly, the dogs are also trained to provide a more public role for the deaf person. Being deaf has been shown to have a major impact on a person's functioning in a 'hearing' society, since people primarily rely on verbal conversation as a means of communication (Kolibiki 2014). Rates of anxiety and depression are higher in deaf people compared with the wider community (Fellinger et al 2005) and rates of emotional problems in deaf children are two times higher than they are for hearing children (Fellinger et al, 2012). Hart et al (1996) found that after being

paired with a hearing dog, the deaf recipient felt safer, less lonely, and less stressed in life. Likewise (Guest et al 2006) found that after obtaining a hearing dog, people felt more included in public life again, had a greater social involvement and an improved feeling of independence. Evidently stemming from the disability being more noticeable to the general public (Hart et al 1996), and seemingly reducing peoples ignorance and improving peoples empathy and understanding towards the disability (Hart et al 1996).

1.2 Justification for thesis

In order to help deaf recipients, the selection of the right dog to become an assistance dog is of fundamental importance to Hearing Dogs For Deaf People, and other organisations that invest time and money in training working dogs. This study seeks to review the reliability of the predictive value of temperament tests currently in use by Hearing Dogs For Deaf People.

Given the crucial benefits an assistance Hearing Dog clearly gives, there has been surprisingly no published attempts to investigate the validity of the behavioural tests used to select the right dog for the job. Indeed, there have been relatively few attempts to investigate assistance dogs' behavioural tests in general. Although most of the attempts that have been made have been on exploring Guide Dog behavioural tests (Batt et al 2008, Mizukoshi et al 2008, Arrata et al, 2011, Tomkins et al 2011, Asher et al 2013), or tests aimed at selecting for multiple types of assistance dogs (Wilsson and Sundgren 1997, Svartberg, K. 2002, 2005, Wilsson and Sinn 2012). However to date there have been none assessing the predictive validity of Hearing Dog's behavioural tests. This thesis seeks to fill this gap, exploring the behavioural tests already in use by the UK based charity Hearing Dogs For Deaf People. In particular it seeks to evaluate whether the current method of testing at 8 months is able to predict performance in the 18 month Hearing dog behavioural test.

Due to the many facets of a Hearing Dog's training life, a further comprehensive review of the tests involving the dogs response to the sounds is beyond the scope of this thesis. Instead this work has purely concentrated on the behavioural requirements of a Hearing dog. As has been discussed there is great importance on a Hearing Dog being able to provide the deaf handler with a feeling of greater inclusion in public life. For this to happen it is of paramount importance that the dog is able to cope and thrive

in a lifestyle that may include going to a plethora of busy public places, while maintaining a friendly attitude towards people, children and other animals. Many of the deaf recipients may also be novel dog owners, so in addition the dogs must be easy for any person to handle and work with in the home, on a walk, or in town environment.

Due to the amount of work that goes into a Hearing Dog, the need therefore to evaluate the predictive validity of the current Hearing Dog behaviour tests is an important one. The work described in the following chapters attempts to thoroughly explore the Hearing Dog tests in use, and then systematically review the quality of the existing published behaviour tests used by other organisations. The predictive validity of the 8 month test (known as test one or T1) currently in use by the Hearing Dogs for Deaf People will then be researched. Convergent validity is evaluated where possible using the CBARQ questionnaire (Serpell and Hsu, 2001) evaluating the dogs behavioural responses to a variety of stimulus over different contexts. A thorough discussion of all results with analysis and recommendations is then provided, in the final chapter.

1.3 The different careers of a Hearing Dog

As a reflection of the different requirements of deaf clients, as well as a reflection of different dogs capabilities, there are three different roles or 'careers' a Hearing Dog may be selected for. The decision as to what career the dog may have is based primarily on two assessments: 'Test one' (T1) at 8 months old and 'Test two' (T2) at 18 months old. The final decision is only ever made after T2. The three different roles are: A full 'Accredited Hearing Dog' (AHD), a 'Sound Support Dog' (SSD) or a 'Confidence and Companion' dog (CCD).

A dog that passes all areas of T1 and T2 and displays a confident relaxed attitude in social and public life, as well as having an aptitude for sound work will become an AHD. These dogs work for the client in their home alerting them to sounds. They will also accompany the client to their place of work, or any public place they go. When in public the dog will wear an burgundy Assistance Hearing Dog coat . This notifies people that the handler is deaf, and along with an ID card signifies the dog

has reached the required health and behavioural standard to be allowed access to all shops and public buildings.

However a dog that displays an aptitude for sound work but fails in any other areas of the test may be selected as an SSD dog. These dog's alert clients to the sounds in the house but are not an accredited Hearing dog, and therefore are not supplied with a Hearing Dog coat or any rights to public access.

Alternatively a dog that is not environmentally confident and does not reliably works to sounds will become a CCD dog. These dogs are given to clients that have a less severe hearing loss, but who still require a dog to provide them with much needed confidence and companionship in the home, and out on walks.

Any dogs that show aggression or that do not fit into any afore mentioned roles may be re-homed to experienced trainers, or if aggression is not apparent to the police as sniffer dogs. Only very occasionally are dogs re homed to volunteers or members of staff, and this is usually only due to health reasons or behavioural problems.

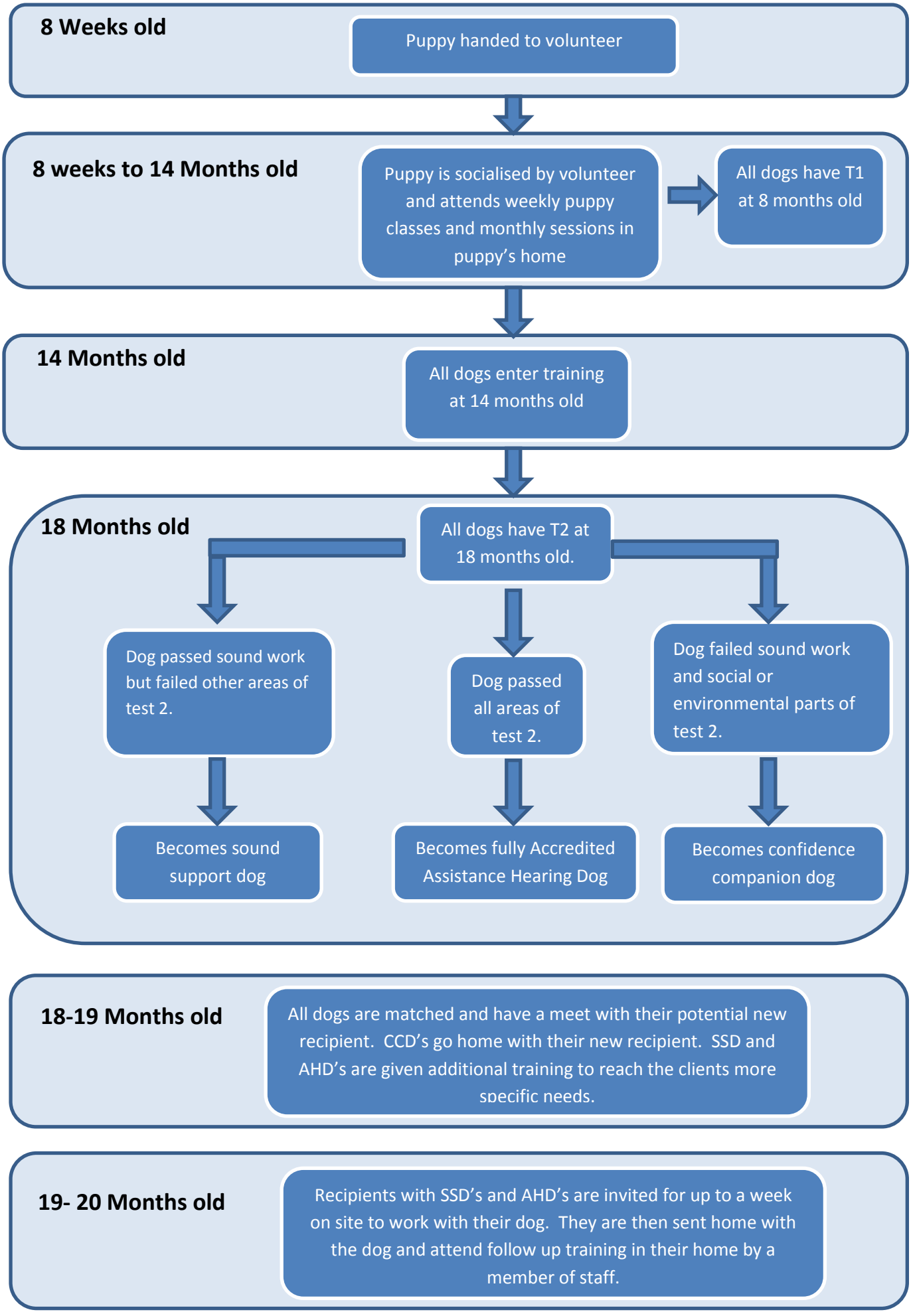
1.4 The life stages Of Hearing Dogs (see fig 1)

At 7 – 8 weeks old a Hearing Dog puppy is handed to a vetted 'puppy socialising' volunteer. The puppy will live with the volunteer until they are approximately 14 months old. During this time the volunteer is responsible for bringing the dog to weekly Hearing Dog puppy classes at a hall by them, run by an experienced member of Hearing Dog staff. Here the dog is taught basic obedience and how to behave around people and other dogs using positive reward based techniques. The volunteer is also required to accept a monthly visit in their home from a member of staff. This is to ensure the dog's health and socialising progress can be checked and recorded. Any additional one to one advice or training is also given during these visits. During this first 14 months of life, the volunteer's primary role is to gradually get the dog used to different types of situations and environments, which the dog may eventually encounter when working as an Assistance Hearing Dog. The dog also goes to stay with other socialisers, and works with other volunteers at least twice during this period, in the hope that they learn to generalise positive experiences to different contexts with different handlers.

At approximately 14 months old the dog usually leaves its socialiser and is relocated to either the Southern (Saunderton, Buckinghamshire) or Northern (Bielby, Yorkshire) Hearing Dog training centres. However if the dog's socialiser already lives close to the centre and is willing to have a trainer work daily within their home, then the dog may instead be trained from that home. If the dog does leave their full time socialiser's home and goes into one of the training centres, then the dog is always placed with a new replacement volunteer who is able to care for the dog every evening and weekend in their home. This practise is to ensure the welfare of the dog and minimise any potential kennel stress on the dogs while in training (Rooney et al 2007).

During advanced training the dog is taught sound work and advanced obedience by an experienced Hearing Dog trainer. Advanced training takes approximately 16 weeks and includes teaching the dog how to respond and alert to sounds, as well as polishing up any obedience needed off lead or on lead. Advanced training also uses the same positive reward based techniques. During this time the dog is also provisionally matched to a deaf recipient, based on the life style the dog can cope with and the personality of both the recipient and the dog. However the recipient is not invited to meet the dog until after the dog has passed his T2 and a final career path has been decided for the dog. (see Fig 1)

Figure 1: Flow diagram to show time line of an Assistance Hearing Dog



1.5 The Assessments of Hearing Dogs

1.5.1 The Quality Assurance Evaluators

The Hearing Dogs' have a formal test at both 8 (T1) and 18 month old (T2). These are carried out by one of four 'Quality Assurance Evaluators' (QAE). One QAE is present at the assessment, accompanied by the handler of the dog. The same QAE may carry out both tests or there may be a different QAE for each test. At each test the QAE has full access to all past behaviour reports and any information that may have been written about the dog from other staff members, therefore they have full access to the dogs past behavioural history. The QAEs realise this may cause potential bias, but in reality due to the real life operational setting, it would be extremely hard to enforce an embargo on all information leaking out to the QAEs. This is partly due to the long period between tests, and partly due to the amount of staff and volunteers involved, working in close proximity.

Three of the QAEs have been dog trainers and quality assurance assessors, each working for Hearing Dogs for over 11 years. The fourth QAE has worked for Hearing Dogs For Deaf People as a QAE for 5 years but has more than 15 years experience training and teaching canine behaviour in other establishments.

The QAEs attend 4 Hearing Dog training courses per year. These span over two days and are either held at the southern Hearing dog training centre in Saunderton, Bucks or the Northern Hearing dog training centre at Bielby, Yorkshire. The training course consists of double blind testing and is filmed by a member of staff. Everybody scoring the dog has very little, if any prior knowledge of the dog. Each evaluator writes down their scores for the assessment separately and without discussion. At the end of the assessment the evaluators hand over their scores which are then later discussed as a group. After a period of 2- 3 months the evaluators re-evaluate the video recordings of the original assessment and score the assessment again. The degree of inter-rater reliability agreement is 81%. The degree of intra-rater reliability ranges from 62% to 85% for the four evaluators.

The intra-rater reliability score is relatively low. This may be due to the technique used to acquire this score. All training assessments are scored on an IPhone camera phone as the assessment is taking place. The picture quality is often poor and shaky, and does not always provide all the information the assessors need to re-score a dog.

An alternative suggestion to this practise could be to use a professionally filmed video clip of a dog's behaviour and body language, and then re-score this clip 2-3 months later.

1.5.2 Location of the assessments

T1 and T2 take place either at the Hearing Dogs Southern centre in Saunderton, Bucks, at their Northern site in Beilby Yorkshire or in the volunteer socialiser's home and local village/small town. The location is chosen depending on proximity of the dog to either site. If the socialiser and the dog is over 1 hour away from either centre, the QAE will carry out the test in the socialiser's local area. This is to ensure that all dogs have a degree of familiarity with the town that they are being tested in, and have been there at least once before. However T2 is primarily conducted at one of the centres, as by this point the dog has entered advanced training as is usually based within at least 30 minutes of either centre.

1.5.3 Test procedures Information from Operational manual in use by Hearing Dogs For Deaf People (Hearing Dogs For Deaf People, 2015)

1.5.3.1 Test 1 (T1)

T1 is performed at approximately 8 months old. The QAE award marks as per the Hearing Dog definitions (See Table 3). Consideration is given to the dog's history from past reports available on the dog, and discussions with the volunteer socialiser and trainer. The dog's behaviour on the day of the assessment is evaluated by the QAE and contributes towards the final score awarded in each category. For each behavioural factor there is also a section where the QAE is required to write notes. These notes should state what the dog was tested with, e.g. for the test of children the QAE should state what age and gender were the child that the dog interacted with, at what location did the dog encounter the child, and what was the child doing. The QAE will also report any relevant history of the dog in this section. Each test takes approximately 3 hours.

The socialiser is either required to bring the dog to their nearest centre at 10.30am or be ready to receive the QAE in their own home at this time. If the assessment is based at one of the centres, the socialiser and dog are greeted by the QAE in reception. Prior to the assessment the socialiser has already been briefed over what the assessment will entail. They have had any questions they may have answered, and been assured that if the dog under performs in any areas on the assessment the dog will still enter training or breeding. Therefore any outcome will have little impact on the socialiser's life with the dog.

The QAE, the socialiser and the dog then travel by car to their nearest park. During the car journey the QAE makes note of how the dogs is affected in the following areas: Vocal reactivity, adaptability (towards novel car), environmental behaviour, recovery rate (if they become anxious) and frustration levels. (see Table 1)

The park is either a football field type recreation area (South), a canal tow path environment (North), or the dog's local park if the assessment is being carried out in the socialiser's local area. The dog is allowed off lead and encouraged to socialise with other dogs, people and children and intermittently play with a toy or a ball if they are available. Hunt and chase levels towards wildlife is observed and recall is also tested at this time. How the dog behaves in the park goes towards their scores in the areas of: Social behaviour adults, social behaviour children, social behaviour dogs, chase, hunt, distractibility, trainability (recall), adaptability (towards novel park), motivation (towards treats or toy/ball, frustration -if not allowed to go see a particular dog, play with a toy), environmental (novel environment) (see Table T1) . If interactions are not recorded in this environment, all effort is made to assess the dog at another time with dogs, people and children. However on rare occasions this is not possible, and in this instance (as with all scoring), previous history is also taken into account.

The QAE, the socialiser and the dog then visit either Princess Risborough (South), Pocklington (North) or a small sized village or town that is local to the socialiser's home. Low population villages/small towns with quiet shops and low volume of 30 mph traffic are chosen, so as to not unduly cause stress to the 8 months old dog. The socialiser is allowed to interact and reward the dog with verbal praise and treats for

correct behaviour around town, however the treats should not be on show the whole time and should be intermittent. Three medium sized shops are visited on the assessment e.g. a chemist, a charity shop, a food shop. People and children are encouraged to say hello and touch the dog. If at any point the dog looks to be not coping in the situation than this is noted and this part of the test is stopped. The dog is then given a score of two or one depending on the severity of the reaction. The town part of the test goes towards the following scoring areas: Vocals, trainability, adaptability, distractibility, social behaviour children, social behaviour dogs, recovery rate, motivation, frustration, environmental behaviour, social behaviour adults. (see Table 1)

While in town the QAE, the socialiser and the dog also frequent a cafe. The dog is allowed a piece of vet bed to lie on. The QAE takes this opportunity to discuss the dogs socialising history further with the socialiser. How the dog behaves in the café part of the test goes towards the following scores: vocals, trainability, distractibility, adaptability, social behaviour children, recovery rate, frustration, environmental behaviour, and social behaviour adults. (see Table 1)

The QAE then drives the socialiser and the dog back to site, or the socialiser's home where the socialiser and the dog have a 15 min break. The dog is allowed another run off lead on site, or in their own garden if the assessment is at the socialiser's home.

After the dog has had a break, in a quiet private room the socialiser is asked to groom the dog in front of the QAE. At the southern centre the room is bare, apart from a sink unit, a table in the middle. A water bowl. Some clean dog bedding and some dog toys. At the Northern centre the room is a single office type environment also equipped with vet bedding, a water bowl and a dog toy. Once in the room the dog is allowed off the lead and is free to explore the room while the QAE and the socialiser chat for approx. 5 minutes. During this time the QAE also notes how the dog is behaving in the room. How the dog behaves in the novel room goes towards the following scores for the tests of: 'Vocals', 'trainability', 'adaptability', 'social behaviour children', 'recovery rate', 'motivation', 'frustration', 'environmental behaviour', and 'social behaviour adults' (see Table 1). If the dog appears overly agitated or attention seeking (barking/jumping up/mouthing) he/she is ignored. They

are interacted with if they choose to pick up and play with one of the toys, or he/she has 4 feet on the ground calmly exploring the room. The dog is never chastised during this time. Once the dog has investigated the room the dog is encouraged to stay still while being groomed. A grooming score is taken, however due to the scope of this paper this is not discussed here. Treats are allowed, however they must be intermittent for good behaviour only and not used to lure the dog or continuously feed him/her. The QAE further takes this opportunity to see how the socialiser and dog interact together.

The socialiser is then requested to leave the room while the QAE works alone with the dog for 15 minutes, initially in the same room and then either on site or along the socialiser's residential road. In the room the QAE will play with the dog off lead with either a soft toy or tennis ball to assess the dog's motivation towards different items, and whether the dog is able to adapt enough to interact with a stranger away from their usual handler. The QAE will then ask the dog basic obedience commands (that are already known to the dog) in the room and then in the outside environment. Treats or a toy are given as a reward at this time, but not used as a lure. The dog is also taught how to target a cooker timer during this time using operant conditioning and a food reward. How the dog behaves when working alone with the QAE goes towards the scores in the following areas: Vocals, trainability, adaptability, frustration, distractibility, recovery rate, motivation, environmental behaviour and social behaviour adults (see Table 1). The dog is then returned to the socialiser

1.5.3.2 Test 2 (T2)

T2 is very similar to T1, however a slightly busier environment is chosen for this test. The exercise area has a moderate population of wildlife instead of a low population, and the town has a slightly higher population and slightly busier traffic. All dogs will have frequented this town at least once a week for the last 10 weeks. Near the Southern centre 'Thame' is used for the town and 'Cuttle brook nature reserve' is used for the exercise area. Near the Northern area 'Beverley' is used for the town, and a different area of the Canal is used that opens up into farms fields. Also On the T2 the QAE does not have a period of working alone with the dog (see Table 2). However the trainer is observed working alone with the dog and handling and grooming the dog, although all of this is no longer in a novel room (see Table 2).

Table 1: *Table to show where each behavioural factor is tested on T1*

Test/Testing environment	Novel testing room	Car travel	Park test	Town test	Settle in cafe	Work alone with QAE
Social Behaviour adults	✓		✓	✓	✓	✓
Social Behaviour children			✓	✓	✓	
Social Behaviour dogs			✓	✓		✓
Environmental	✓	✓	✓	✓	✓	
Recovery rate	✓	✓		✓	✓	✓
Adaptability	✓	✓	✓	✓	✓	✓
Vocals	✓	✓	✓	✓	✓	✓
Motivation	✓		✓			✓
Trainability	✓					✓
Frustration	✓		✓	✓	✓	✓
Chase			✓			
Hunt			✓			
Distractibility	✓		✓	✓	✓	✓

Table 2: *Table to show where each behavioural factor is tested on T2*

Test/Testing environment	Novel testing room	Car travel	Park test	Town test	Settle in cafe	Work alone with QAE
Social Behaviour adults			✓	✓	✓	
Social Behaviour children			✓	✓	✓	
Social Behaviour dogs			✓	✓		
Environmental		✓	✓	✓	✓	
Recovery rate		✓		✓	✓	
Adaptability		✓	✓	✓	✓	
Vocals		✓	✓	✓	✓	
Motivation			✓			
Trainability						
Frustration			✓	✓	✓	
Chase			✓			
Hunt			✓			
Distractibility			✓	✓	✓	

1.5.4 Scoring criteria for Hearing Dogs

The first battery of tests at 8 months (T1) and the second battery of tests at18 months (T2) are all scored on a scale, where 1 represents very poor, to 5 which represents very good. A pass is a 3, 4 or 5 grade, where as a fail is a 1 or 2 grade. All dog must score a 3 or above in all areas to have passed the overall test. However any dog who fails any or all areas of T1 will not be disqualified from taking T2. Any dog that fails any areas of T2 will then be discussed by the Hearing Dog senior management team. If there are mitigating circumstances why the dog may have passed on T1 but failed on T2, then the dog may be re-tested at a period of 2 to 4 weeks later. However not every dog is re-tested. It is done on a case by case base, and the decision relies strongly on the dog's history. If the dog fails the re-test then the dog will have a career change or be re homed. If a dog failed its T1 and T2 in any area and there are no mitigating circumstances then a further re test is not given and a career change, or possible re-homing is decided upon.

1.5.5. The different behavioural factors (Information from existing manual- Hearing Dogs For Deaf People, 2015)

1.5.5.1 Social behaviour (adults) (see Table 3)

This area is measured on the dog's behaviour when stroked on the head and body by both men and women. This is tested off lead in the Park and novel testing room, and on lead in town and settle in café parts of the assessment. However on the T2 it is no longer tested in the novel testing room. As with all tests consideration is also given to the dog's history from past behaviour reports, however if people are not encountered naturally in these environments then they are re-tested with different age and sex stooges at a later time, either at one of the centres or in the dogs local area.

1.5.5.2 Social behaviour (Children)(see Table 3)

This area is measured on a dog's behaviour when being stroked on the head and body by children under the age of ten. This is usually observed either in the park, town and settle in café. If children are not seen at this time, then every effort is made to see them either later that day or at another time within a month's time. However consideration is also given to the dog's history.

1.5.5.3 Social behaviour (dogs) (See Table 3)

This is measured on a dog's behaviour when interacting off lead and on lead with other dogs. Every consideration is made to test the dog with both sexes of dog. Big and small dogs, and old and young dogs. This is tested off lead in the Park and on lead in town in bothT1 and T2. If dogs have not been encountered naturally by the end of each test, then the dogs are re tested on another day either naturally or with different stooge dogs within a months' time. History is also taken into account.

1.5.5.4 Environmental Behaviour (see Table 3)

Environmental behaviour is measured on a dog's confidence in a variety of different environments (including public transport). As per previous tests consideration is also given to the dog's history, from past behaviour reports carried out by any members of staff who have worked with the dog, and from discussions with the volunteer socialiser. 'Environmental behaviour' is tested in the following areas of the T1: car travel, Park, town, novel testing room, settle in café on the 8 month test. It is mostly tested in the same areas on T2, however with the exclusion of the novel testing room. The dog is walked around town for a period of 45 minutes to one hour. During this time they frequent at least three different shops, including one food shop. If at any point the dog shows extreme anxiety the test is stopped and the dog is returned back home.

1.5.5.5 Recovery rate (see Table 3)

'Recovery rate' is measured by how quickly the dog returns to a normal acceptable state after a setback or knock. To score this the dog is not exposed to any organised event/exercise that is predicted to cause the dog a knock/setback. The score is based on the dog's history and how they behave at the time of the test. If the dog does not have any known contexts that cause them a setback it is likely they will be scoring high. For the 8 month test recovery rate is tested when in car travel, town test, settle in café, novel testing room, working alone with QAE . However it is only scored in the areas of car travel, town test, settle in café in the 18 month test.

1.5.5.6 Adaptability (see Table 3)

'Adaptability' is scored on the dog's ability to adjust to changes in their routine, handler and environment. Adaptability is measured by comparing the dog's behaviour

when these factors are constant to when they are altered. For the purposes of this test, a dip in environmental or social confidence after transferring from socialising to training could be a symptom of poor adaptability. On T1 adaptability is measured by Novel testing room and when working alone with QAE. On T2 adaptability is measured by the dogs past history and in the areas of; car travel, park test, town test, settle in café. On the T2 the areas of novel testing room and working alone with the QAE are not used to score this area.

1.5.5.7 Vocal Reactivity (see Table 3)

Vocal reactivity is measured by the quantity of vocal reactions to events and situations. This is based on recent history in socialising and the scores on the behaviour test. It is tested in: car travel, park test, town test, settle in café, Novel testing room on the 8 month test (see Fig 7). It is tested in all of the same areas in T2 with the exception of the novel testing room.

1.5.5.8 Motivation (see Table 3)

'Motivation' is measured by the value & quantity of reinforcement required for a known command. On T1 the tester plays with the dog on the floor for a period of 10 minutes with different items, the tester then asks the dogs to complete known commands for different items of reward. If the dog is hard to engage or does not engage in play or commands for any item then the dog would score low in this test. On T2 the scoring for motivation is based on recent history and the dog's behaviour when responding to play and known commands for the trainer. If the dog is hard to engage or does not engage in play or commands for any item for the trainer then the dog would score low in this test On the 8 month test motivation is tested in the following areas; park test, novel testing room and when working alone with the QAE. On T2 the dog is not tested in the novel room, or working alone with the QAE.

1.5.5.9 Trainability (see Table 3)

'Trainability' is measured on the dog's ability to learn new behaviour/s. The score definition is based on operant conditioning and targeting for positive reinforcement. The score for Trainability takes into account recent history, and from observing the dog in the following areas in the behaviour test : novel testing room, work alone with QAE . On T1 the dog is taught to target a cooker timer. On T2 the areas of novel

testing room and working alone with the QAE are not used to score trainability, instead the scores are based on the dog's recent ability to learn advanced sound work.

1.5.5.10 Frustration (see Table 3)

Frustration is measured by the ability to cope when something of value is withheld or removed. At 8 months this is scored by observing the dogs in the town, the park, the café, the novel testing room and when working alone with the QAE . On T2 this is not observed in the novel testing room and the dog does not have a period of time working alone with the QAE.

1.5.5.11 Chase (see Table 3)

Chase is measured by the dog's motivation to chase moving objects. This is tested in the park for bothT1 and T2 as well as taking the recent history of the dog into consideration.

1.5.5.12 Hunt (see Table 3)

Hunt is measured by the dog's motivation to hunt for wildlife. This is tested in the park for bothT1 and T2 as well as taking the recent history of the dog into consideration.

1.5.5.13 Distractibility (see Table 3)

Distractibility is measured by the amount of interest the dog is showing in environmental stimuli. On T1 it is scored by recent history and by observing the dog on the park test, the town test, the settle in the café, in the novel testing room and when working alone with the QAE. On T2 it is also scored by history and by observing the dog on the park test, the town test and the settle in the café.

Table 3: *Table to show the Hearing Dogs For Deaf People Scoring definitions for T1 and T2*

Test/Score Definition	5	4	3	2	1
Social Behaviour adults	Dog is confident to meet people.	Dog is generally confident to meet people, however has occasionally shown a sensitive side in unusual situations, good recovery.	Dog is generally confident with people. Dog may have some sensitivities however is not showing any undesirable behaviour	Dog is lacking confidence with people in one or both areas. Dog is showing undesirable behaviour level 1	Dog is lacking confidence socially in one or multiple areas. Dog is showing undesirable behaviour level 2
Social Behaviour Children	Dog is confident to meet children.	Dog is generally confident to meet children, however has occasionally shown a sensitive side in unusual situations, good recovery.	Dog is generally confident with children. Dog may have some sensitivities however is not showing any undesirable behaviour	Dog is lacking confidence with children in one or both areas. Dog is showing undesirable behaviour level 1	Dog is lacking confidence with children in one or multiple areas. Dog is showing undesirable behaviour level 2
Social Behaviour Dogs	Confident with other dogs and interacts appropriately	Confident with other dogs. Does have a history of sensitivities/OTT behaviour, however now interacts appropriately	Generally confident with other dogs. May be sensitive/some OTT behaviours however no undesirable behaviour seen.	Dog is lacking confident with other dogs and/or is showing undesirable behaviour level 1	Dog is lacking confidence with other dogs and/or is showing undesirable behaviour level 2
Environmental	Dog is confident in all environments	Dog is generally confident in all environments. However has occasionally shown a sensitive side in unusual environments, good recovery.	Acceptable level of confidence however has shown a sensitive side in some environments/ situations, good recovery.	Dog is lacking confidence in one or multiple environments. Dog is showing undesirable behaviour level 1	Dog is lacking confidence environmentally. Dog is showing undesirable behaviour level 2.

Test/Score Definition	5	4	3	2	1
Adaptability	No change to the dog's behaviour when routine, handler or environment changes	No significant change to dog's behaviour when routine or environment changes although takes a short while to adjust to changes in handler	No significant change to the dog's behaviour when routine changes although takes a while to adjust to changes in handler or environment. Has a history of showing some level 1 undesirable behaviour but these improve to an acceptable level once dog has adjusted.	Significant changes to dog's behaviour (I.E unexpected or level 2 undesirable behaviours seen) when routine/handler or environment changes	Takes a long time/or does not adjust to changes in routine, handler or environment
Vocals	No vocal reaction seen in any situations/environments	Low level vocal reaction to novel stimuli	Occasional vocal reaction to stimuli however, good recovery	Frequent vocal reactions to some stimuli with low recovery	Frequent or extreme vocal reactions in most situations/environments with low or no recovery
Motivation	Motivated by social interaction and low level reward	Motivated by social interaction and/or intermediate level of reward	Motivated by social interaction and/or high level reward	Lacking motivation, though occasionally shows interest in high value reward	Difficult/unable to motivate

Test/Score Definition	5	4	3	2	1
Frustration	Calm and relaxed with no frustration behaviours seen	Low level frustration behaviours seen with good recovery	Occasional frustration behaviours seen with good recovery	Frequent frustration behaviours seen with slow recovery	Frequent or extreme frustration behaviour seen I.E continuous mouthing/vocal with slow or no recovery
Chase	No/Low motivation to chase	Opportunistic chase seen however quickly loses interest	Chase seen (Dog must achieve a recall score of 4 or above)	Chase seen frequently with high arousal levels (dog would achieve a recall score of 3 or below)	High arousal predatory chase
Hunt	No/Low motivation to hunt	Hunt drive seen however quickly loses interest	Hunt drive seen (dog must achieve a recall score of 4 or above)	Hunt drive seen frequently with high arousal levels (dog would achieve a recall score of 3 or below)	Strong hunt drive with high arousal
Distractibility	No/low interest in environmental stimuli	Low/medium interest in environmental stimuli however quickly loses interest	Interest in environmental stimuli however not excessive (dog must achieve a recall score of 4 or above)	Frequent interest in environmental stimuli with high arousal levels (dog would achieve a recall score of 4 or below)	Strong interest in environmental stimuli with high arousal levels

1.5.6 Initial observations of the Hearing Dog Tests.

From initial observations, possible strengths of the Hearing Dog tests may be the method of naturalistic observation in use. The dogs are tested in the type of environment in which they will eventually work, as opposed to an artificial laboratory setting. This may have the advantage of allowing a better understanding of how the dog will eventually cope in such a setting; however it has the disadvantage of the researcher having less control over the variables and the ability to be able to make firm conclusions of cause or affect.

A dog's emotion and affective state may influence the dog's cognitive abilities, and have a knock on effect when it comes to tests such as the trainability test, or even how the dog is able to respond to people, dogs or children in that environment. The

organisation is aware of this and therefore takes the dog's history in to account when grading all areas. Assessments of Hearing Dogs may benefit from this knowledge and sensitivity of the dogs prior behaviour, however in turn this may also cause potential bias towards the observer's interpretation of the dog's behaviour.

Importantly the Hearing Dog tests, give consideration to the dog's reactions towards social factors such as people and dogs. This is explored in such a way that consideration is given to how the dog reacts to different sexes and ages. However the same consideration is not given when assessing a dog with a child. The current 'child' test does not stipulate what sex the child should be, and the wide age range from baby to 10 year old may greatly influence different reactions across tests.

In the majority of the Hearing Dog tests the dog's familiar handler is present. A dog's relationship with their handler and its ability to interact with their handler may have an effect on the assessment situation. This could affect many of the test components including trainability. This may be a negative or a positive to the Hearing Dog tests. One factor which also bears further consideration is that between T1 and T2 the dog also has a change of handler from familiar socialiser to trainer.

Importantly the operational details of T1 and T2 differ slightly from each other in three ways: A) In T2 the novel testing room is no longer used. B) In T2 the QAE no longer has a period of working alone with the dog C) A slightly busier town or park environment may be used in T2 compared with test T1. Hearing Dogs For Deaf People justifies point A and B by stating that for A and B the same scenarios can no longer be novel. By T2 the QAE cannot grantee that any areas or rooms on site that haven't been frequented by the dog already. Likewise there is a good chance that the dog would have already worked for the QAE before on their previous test, therefore this is also no longer a novel situation in which to test the dog's adaptability and recovery rate by.

Hearing Dogs justify point C, by explaining that the dog may have become over used to the quiet town and park. By simply replicating test 1 the charity feels it may no longer be a true reflection of how the dog may cope when placed with its recipient working in a novel or slightly busier area. These inconsistent testing conditions may of course result in 'measurement error' and may cause the individual's score to vary significantly from test 1 to test 2.

Between the ages of 8 to 18 months a dog goes through a period of adolescence. Test one is taken in the middle of this period and the dog's response may therefore be affected by his or her hormonal state. Likewise by test two the dog has been neutered. Of course, it may be unlikely the exact same results will be obtained on each test, but a strong positive correlation between the results of the tests, when the dog's has differing sexual states, may indicate behavioural consistency and reliability over time.

One of the main benefits of the existing Hearing Dog Tests may be that every dog has a T2 irrespective of a pass or failure on T1. This allows a further analysis of the reliability of the predictive value of temperament tests currently in use by Hearing Dogs For Deaf People.

The use of tests to provide an efficient filter for quality of dog as an effective method has the advantage of potentially being cost efficient for the charity, however until the findings are fully analysed, staff members need to be cautious about using such tests until there is more evidence to support any findings. This thesis will critically analyse existing published material over the following chapters as well as analyse the reliability of the predictive value of the operational temperament tests currently being carried out on assistance hearing dogs.

Chapter 2

A systematic review of the published predictive validity of assessments with respect to task performance in dogs.

Contents **Page**

A systematic review of the published predictive validity of assessments with respect to task performance in dogs.

2.1 Introduction

Significant focus has been applied to creating temperament, and selection tests that aim to predict performance in working dogs (Murphey 1995; Seksel et al 1999; Slabbert and Odendaal 1999; Batt et al 2008; Mizukoshi 2008; Sforzini et al 2009; Sinn et al 2010; Wilsson and Sinn 2012; Asher et al 2013). A selection test can involve scoring the reaction of dogs against a series of controlled stimuli and using this information to then make predictions about working aptitude (Asher et al, 2013). Selection tests can be observed either in the dog's usual environment or in contrast when they are in standardised conditions of a laboratory. Many selection tests are based on behaviour or temperament tests. Serpell and Hsu, 2001, define animal behaviour tests as standardised experimental situations where stimuli serve to elicit behaviour that is statistically compared with that of other individuals placed in the same situations, in order to classify the subject tested. These are not the same as temperament tests. Temperament tests look for differences in an individual's behaviour that are consistently displayed when tested under similar situations (Diedrich and Giffoy 2006).

Selection and behaviour tests are often generalised and related to how the dog is likely to perform in real life working situations. By performance it is meant *"A task or operation seen in terms of how successfully it is performed" (Oxford Dictionary 2014)*. In other words to have predictive validity the initial assessment must be measured against the success of the working task in the field, which must be measurable. Diederich and Giffroy (2009) highlight that in order to attain predictive validity, results from behavioural tests during the animals development must be correlated with the degree of success in later performance. Based on this, this review questions how many of the published assessments measure working performance as opposed to a later test or some other proxy of performance.

The original aim of this review was to explore the predictive validity of the published assessments with regards to working ability. Working ability is defined by Assistance dogs international (2015) as *"a consistent ability to complete a minimum of three*

operational tasks for its handler". A working dog in other words must be able to perform a functional role for an organisation and perform set tasks above just giving companionship. However due to the limited published research on just working dogs this review has been expanded to include dogs that undergo an assessment for pet roles also

This systematic review analyses and evaluates published assessments of temperament and behaviour tests that claim to predict performance in dogs.

2.2 Method

This review followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. PRISMA is recommended set of requirements when reporting a systematic reviews or meta-analyses. It is the basis for reporting style in this systematic review.

2.2.1 Eligibility criteria

2.2.1.1 Age of dogs

This review included published research of dogs at any age. The rationale for this was based on the evidence that due to the different roles, published assessment tests were commonly conducted at different ages depending on the life the dog was being assessed for.

2.2.1.2 Sex of dogs

Working dog programmes rarely (if ever) exclude dogs due to their sex. Consequently this systematic review also included published papers on dogs of any sex.

2.2.1.3 Follow up period

The validity of the behavioural tests and the stages that the various assessments took place were examined in order to look at how, and at what threshold task performance was being graded at. This was done by reviewing the tests and the follow up procedures, if any, that were being used. In particular this review was interested in how many (if any) also followed up on the dogs that did not graduate.

2.2.2. Information sources and search strategy

This systematic review used electronic peer reviewed published papers from Science Direct, Scopus and Google scholar. The search terms of 'dog personality', 'working dog' 'temperament tests', 'working dog behavioural test', 'working dog assessments', 'assistance dogs', 'dog temperament tests', 'dog behavioural test', 'Assistance dog temperament', 'predictability of dog behaviour', 'dog personality', 'dog performance', 'dog behavioural test', 'service dog temperament tests' and 'service dog selection tests' were searched for across all three academic data bases, during the period of 20/10/14 to 08/11/14.

2.3 Results and Discussion

Specific results of the study selection and characteristics with regards to task performance and will be discussed further in this section. This will be followed by a critique of the different types of behaviour test typically used.

2.3.1. Study selection

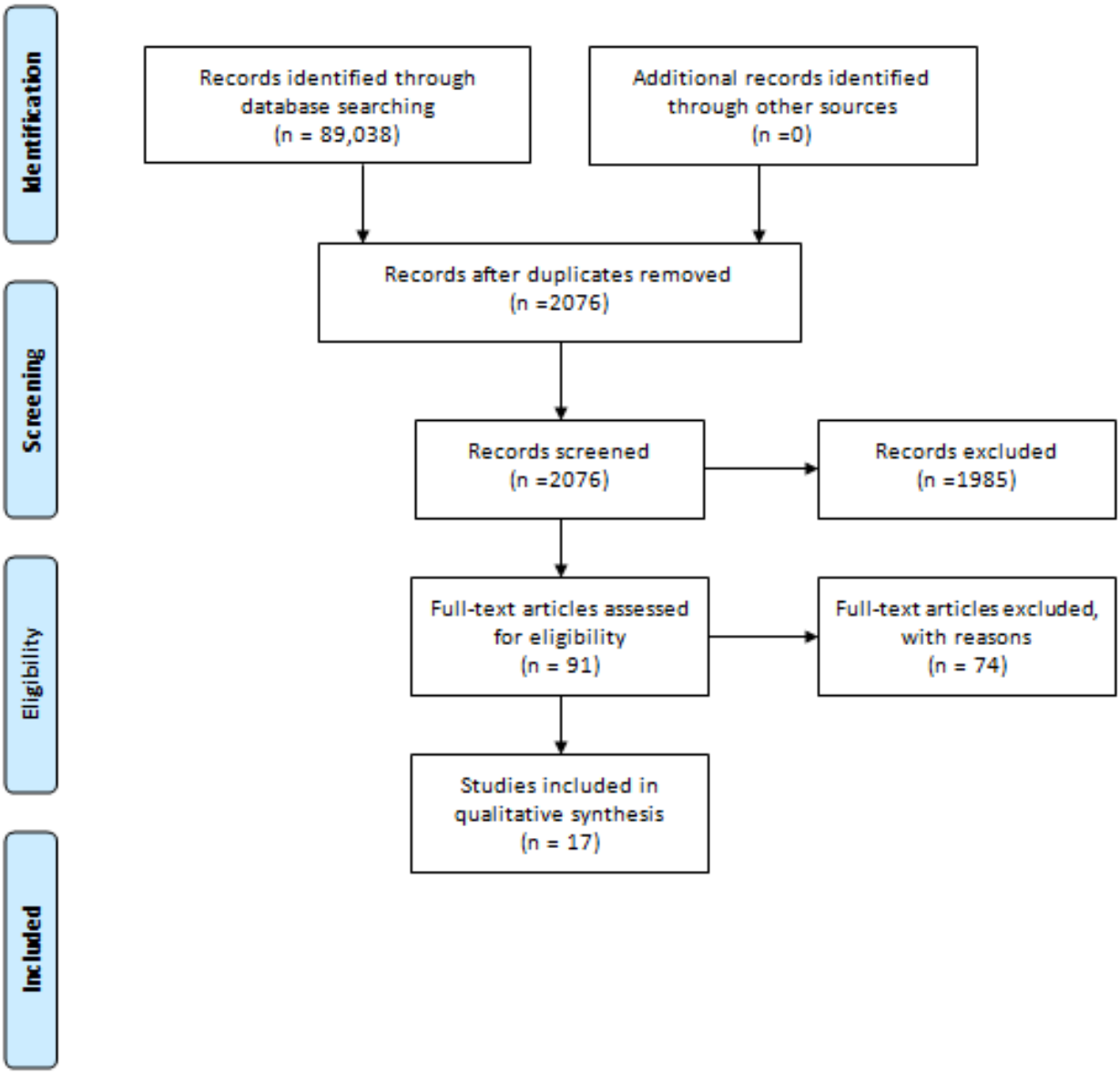
A total of 89, 038 published papers were initially found using the previously discussed search terms. 86, 962 of these were deemed duplicates. After applying a filter of 'dog' to the remainder 2076 papers, 1093 papers were highlighted. 179 of these were further removed due to being book excerpts, leaving 914. Any published papers purely exploring anatomy or health were removed leaving 551 papers. Abstracts of the 551 papers were further examined and deemed not relevant for this thesis if they did not discuss behaviour, leaving 91 papers further assessed for their eligibility.

91 studies were narrowed down, they were further screened for eligibility. Studies without behavioural assessments and just containing questionnaires were excluded due to the scope of this review. However studies that contained questionnaires alongside assessments were included. Temperament or behaviour tests were considered of value for this review. Of the 91 papers, 20 were further excluded due to the lack of a temperament or behaviour test.

This Left 71 papers in the systematic review. 10 of these were review papers and had insufficient information on how data was obtained and so were excluded. Leaving 61

papers. 44 were not attempting to look at performance and were primarily concerned with behavioural tests concerning management of dogs or stress of dogs. Leaving 17 papers: Slabbert and Odendaal 1999; Svartberg 2002; Lucidi et al 2005; Svartberg 2005; Christensen et al 2007; Batt et al 2008, Bollen and Horowitz 2008; Mizukoshi et al 2008; Vas et al 2008; Seksel 2009; Sforzini et al 2009; Sinn 2010; Asher et al 2011; Tomkins et all 2011; Valsecchi et al 2011; Barnard et al 2012; Wilsson and Sinn 2012. that were included in this review of predictive validity of real task performance. *(see figure 2)*

Figure 2: *PRISMA Flow chart to show screening and eligibility process for this systematic view*



2.3.2. Study characteristics (Task performance) (see Table 4)

Originally this review was primarily concerned with which studies were able to demonstrate predictive validity with respect to final task. In order for a study to do this it was considered that the assessments should test performance of the task in real

life. However as only Valsecchi et al (2001) and Vas et al (2008) carried out behavioural testing past graduation in a real life setting including both the past and failed dogs, it was necessary to explore the various behaviour tests that occurred prior to the dog graduating into a working or pet home.

Only by re-testing dogs' that have both passed and failed can performance and predictive validity of a test be measured. Out of the 17 papers, only 6 conducted behaviour re-tests on both passed and failed dogs. These were: Slabbert and Odendaal 1999; Seksel et al 1999; Batt et al 2008; Mizukoshi et al 2008; Vas et al 2008; Valsecchi et al 2011.

Out of the 6 studies that conducted re- tests on both passed and failed dogs only Seksel et al (1999), Valsecchi et al (2001) and Vas et al (2008) followed the dog into placement/real life. Seksel et al (1999) did this in the form of a questionnaires and individual personality judgements; however Valsecchi et al (2001) and Vas et al (2008) performed a behaviour re- tests by direct observation in the dog's natural environment.

Vas et al (2008) were looking at the consistency of a dogs' performance and behaviour during different approaches from people. Their sample group consisted of observing pet dogs' behaviour in a familiar park, so the test conditions did not change over time. They found that the dog's performance in the tests were consistent over time and therefore the first test was able to show predictive validity.

Conversely Valsecchi et al (2001) were looking at the behavioural responses of shelter dogs' pre and post placement. Their first two tests were conducted in the behaviour centre however a further one was conducted once placed in a new pet home. They found that the dogs correlation of scores and overall the temperament in the shelter remained similar to what the dog later expressed at home. By conducting re- tests on both passed and failed dogs in a natural setting, their first tests were able to show predictive validity with respect to task performance in a real life setting.

From the rest of the papers in this review, Sforzini et al (2009), Sinn et al (2010), Tomkins et al (2011) and Wilsson and Sinn (2012) only carried out re-tests on dogs that passed and not on any dogs that failed.

Out of the 17 papers, 7 papers (Svartberg 2002; Svartberg 2005, Lucidi et al 2005; Christensen et al 2007; Bollen and Horowitz 2008; Asher et al 2011; Barnard et al 2012) did not do any behaviour re-test of any kind.

Interestingly Batt et al (2008) also gave a behaviour test to graduating and non-graduating guide dogs, however the non-graduating dogs were never placed as working guide dogs, so again true working performance could not be validated.

Two papers (Christensen et al 2007; Bollen and Horowitz 2008) did conduct a follow up phone call, but only on the dogs that passed graduation, meaning the dogs that failed were never monitored to see how they could have performed in real life. One paper (Lucidi et al 2005) conducted a follow up review 1 year after placement, but again, only on the dogs that had passed graduation.

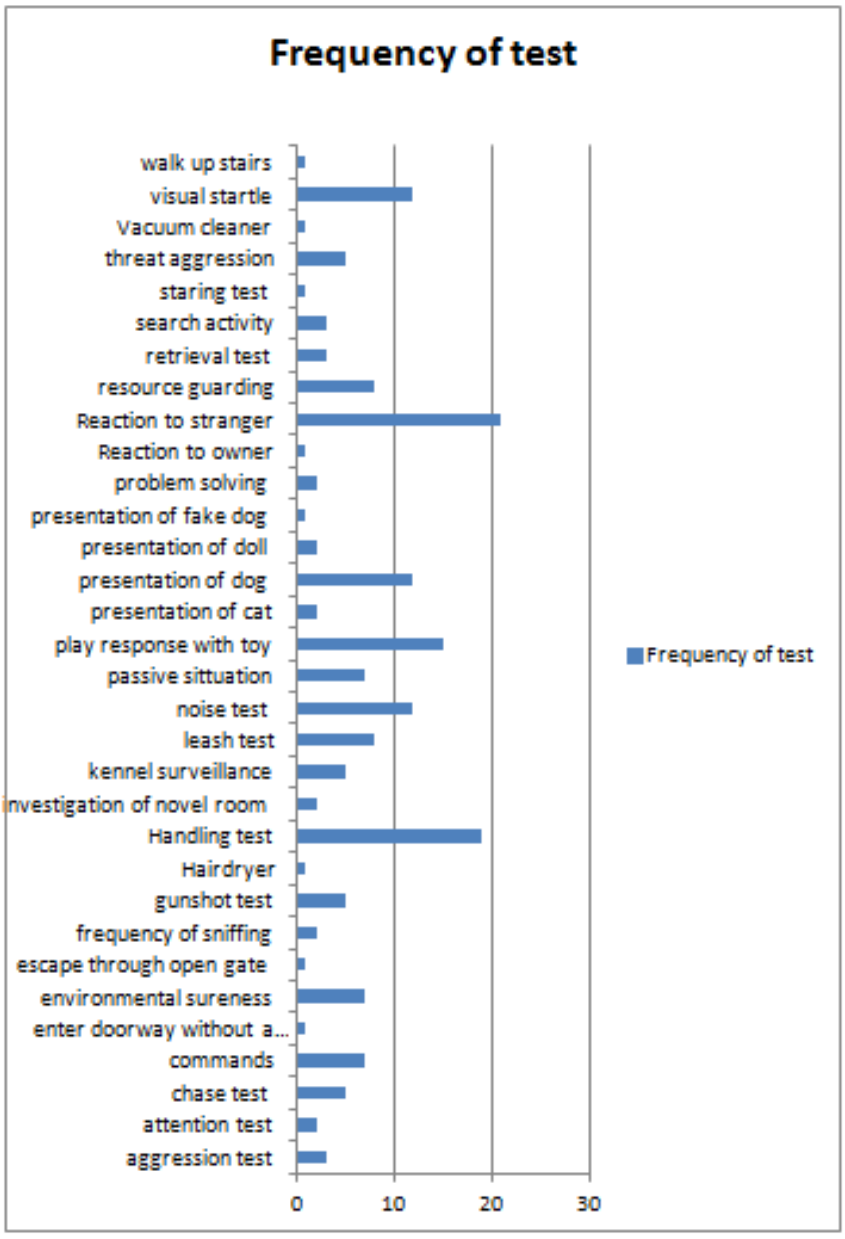
***Table 4**: Follow up period of the behaviour tests reported in the 17 scientific papers*

Reference	Behaviour test 2	Re-test given to dogs who both passed and failed	Follow up method past graduation of behaviour test.	Both graduating and non-graduating dogs monitored into placement
Barnard et al (2012)	1 x behaviour test on 34 volunteer pet dogs	N/A	Intake Questionnaire before behaviour test. No follow up after	No
Bollen & Horowitz (2008)	1 x behavioural test on 2017 shelter dogs	N/A	Phone call 6 months after placement.	No
Christensen et al (2007)	1 x behaviour test on 279 dogs	N/A	Phone call 13 months after placement.	No
Seksel, Mazurski, and Taylor (1999)	Behaviour tests repeated x 3 on 60 dogs	Yes	Questionnaire completed 6 months after behaviour tests.	yes
Svartberg, (2005)	1 x Swedish 'DMA' behaviour test on 697 dogs	N/A	Questionnaire completed 1-2 yrs later.	Yes
Vas et al (2008)	Behaviour tests repeated x 2 on 49 dogs	Yes	Retest done, immediately, 6 months or 1 year later in familiar park.	Yes
Valsecchi et al (2011)	Behaviour tests repeated x 2 on shelter dogs.	Yes	Follow up behaviour test 4 months after adoption.	Yes
Asher et al (2013)	1 x behaviour test on 587 puppies.	N/A	General success in Guide dog training was monitored	No
Mizukoshi et al (2008)	Behaviour tests every 2 weeks for 12 weeks on 8 dogs	Yes	Success rate based on dogs ability to pass final examination. No follow up	No
Batt et al (2008)	behavior tests repeated x 3 on 43 dogs	Yes	Behaviour retests taken at 6, 14 and 20 months.	No
Lucidi et al (2005)	1 x behaviour test on 23 shelter dogs	N/A	Follow up review 1 yr after dogs have been out working	No
Sforzini et al (2009)	Behaviour tests at 5, 7 and 9 months of age on 32 dogs	No	Behaviour test given 15 months later.	No
Slabbert and Odendaal (1999)	Behaviour tests at 8 wks, 12 & 16 wks, 6 and 9 mnths on 167 dogs	Yes	Success rate based on dogs ability to pass final examination. No follow up	No
Sinn, Gosling and Hilliard. (2010)	Behavioural test repeated x 2.	No	Success rate based on dogs ability to pass final examination. No follow up	No
Svartberg (2002)	1 x DMA behaviour tests on 15,329 dogs.	No	No follow up	No
Tomkins, Thomson, and McGreevy (2011)	1 x behaviour test on 113 dogs	No	Success rate based on number of dogs that graduated. No follow up	No
Wilsson and Sinn (2012)	1 x behaviour tests on 200 armed forces dogs.	No	Success rate based on number of dogs that graduating. No follow up	No

2.3.3 Types of behaviour test

The 17 papers in this review were further screened for the type of activity in the behaviour tests that were being used. (See Table 5). Many of the papers have multiple tests seemingly testing the same thing. Due to the scope of this paper the frequency of these have been recorded and have been grouped where necessary under one heading. The methodology and predictive validity of the most common behaviour tests will be discussed further in the results.

Table 5: Bar chart to show the type and frequency of tests within the Peer reviewed journals in this review



2.3.4 Resource guarding

Resource guarding can be defined as human or dog directed aggression because of possessive behaviour over food or objects (Serpel and Hsu 2001). The studies claiming to test for this were reviewed.

Five of the papers in this systematic review (Christensen et al 2007; Bollen and Horowitz. 2008; Sforzini et al. 2009; Sinn et al. 2010; Valsecchi et al. 2011) purport to test for the trait of possessiveness in their subjects. Content validity, a form of face validity refers to how well a test appears to measure the behaviour for which it is intended. For a test to have face validity it must be able to measures the specific character traits intended, and importantly limit any confounding variables that may skew the measurement. However all of the tests used to assess possessiveness described above are conducted in a novel room or field, and all except for Sforzini et al (2009) has a novel unfamiliar handler testing the dog. Tests conducted in this manner will be subject to the influence of the novel environment and presence of the stranger, and neophobia in this context will obviously impact on the ability to assess possessiveness. The way the subject reacts in this situation may be completely different to how they behave once settled in a familiar safe environment with a trusted person.

Bollen and Horrowitz (2009) and Christensen et al (2007) also use a plastic or rubber hand to reach inside the dog's bowl. This presents a further confounding variable. Although safety when working with unknown rescue dogs needs to be taken into consideration, a plastic hand on a stick is not representative of a real hand; it is devoid of the familiar smell and absence of relationship with the real hand. It seems reasonable to suggest that this may be simply be a novel object to a dog, perhaps a toy to be played with or a new object to be suspicious of. How they behave towards a rubber hand on a stick compared to real life situations, is questionable (Tempany and Mills 2008).

Another confounding variable which should always be taken into account when conducting any tests such as these is the subject of motivation. Christensen et al

(2007), Bollen and Horowitz (2008) and Sforzini et al (2009), all claim that their test using food bowls is able to show predictive validity for possession reactivity in real life. But it could be questioned whether it is predicting possession in general or just food possession. They have not first assessed whether the food is of value to each of their subjects. It is remiss to say that every dog values the same type of food at all times of the day and in every situation, just as it would be to say that of humans. It is also not valid to suggest that tests only using food are representative of all types of possession in general. One dog when threatened or insecure may feel the need to guard a familiar human or favourite bed, but the same dog may be perfectly fine when a stranger puts their hand in their bowl. Generalisation over something that could potential lead to a dog becoming distressed or at worst biting are risky and should be avoided.

2.3.5 Presentation of a dog

For 'presentation of a dog' any behaviour test aiming to test for how the subject reacted to a conspecific was reviewed. Eight of the papers in this systematic review (Seksel et al 1999; Lucidi et al 2005; Christensen et al 2007; Batt et al 2008; Bollen and Horowitz 2008; Tomkins et al 2011; Valsecchi et al 2011; Barnard et al 2012) introduce their subject to either another dog or a fake dog to ascertain how the subject behaves in the presence of other dogs. The authors purport their tests measures either dog to dog interaction (Barnard et al 2012; Bollen & Horowitz 2008) intraspecific aggression/sociability (Christensen et al 2007; Valsecchi et al 2011) dogs' aptitude to meeting other dogs (lucidi et al 2005) or dog distraction (Tomkins et al 2011). However only Valsecchi et al (2011) and Batt et al (2008) have commented which sex the conspecific were and only Batt et al (2008) has commented on the breed and sexual status of the conspecific.

All these tests aim to improve our knowledge of the individual dog, however it is questionable if it is possible to generalise to everyday behaviour after only testing with such a limited type of conspecific. For example, (Goddard & Beilharz, 1985) found that potential Guide Dogs reacted differently when presented with a juvenile compared to an adult male conspecific. Whereas Fatjo et al (2007) when looking at captive wolves found that most agonistic interactions were between animals of the

same gender and particularly between males. Christensen et al. (2007), found that once their 67 successful test subjects were re-homed, intra specific aggression in particular were not predicted and detected reliably. This was perhaps because their test was only done on one occasion with one 'friendly' conspecific in a novel room.

Misinterpretation of motive could also be problematic in this subject area, agonistic behaviours can be a way of avoiding the escalation of conflict and aggression. For example ambivalent signals such as showing teeth could be a strategy to avoid the escalation of aggression and could be used as a useful balance between cohesion and conflict (Packard, 2003). Furthermore, the meaning of aggression could vary depending on whether it is linked to active or passive submission in the dog (Sueda and Malamed 2014) that many of the tests do not refer to.

2.3.6 Reaction to a stranger

Nine papers (Seksel et al 1999; Lucidi et al 2005; Svartberg et al 2002; Svartberg et al 2005; Christensen et al 2007; batt et al 2008; Bollen and Horowitz 2008; Vas et al 2008; Barnard et al 2012) explore how their subjects respond to the approach of a stranger. Ratcliffe, McComb, and Reby (2014) found that dogs who had prior experience of both men and women could categorize human gender and orientate themselves correctly towards a stranger after hearing a tape recording of a male or female voice, whereas Wells and Hepper (1999) found that in their subject group there was a stronger decrease in barking and tendency to look towards the human whenever the subject was a woman compared to a man. However in this review only Seksel et al (1999) assess their dog's reactions to both male and female strangers. The other authors either do not refer to what sex the stranger was (Svartberg, 2002, 2005, Lucidi et al 2005, Christensen et al 2007, Batt et al 2008, Valsecchi et al 2011) or have tested with only a female (Bollen and Horowitz 2008, Vas et al 2008, Barnard et al 2012). Thus prior experience could propagate a negative bias towards one sex or the other so unless the dogs are tested with both, it is questionable whether they show predictive validity with respect to task performance

Another confounding variable comes from whether the dog was leashed or not since this may induce frustration. From the 9 papers exploring how subjects react to a stranger, Lucida et al 2005, Vas et al 2008 and Barnard et al 2012 all had their subjects leashed while testing with a stranger. Batt et al (2008) and Barnard et al (2012) also conducted their test within a novel room. Significantly Ortolani et al (2009) found Ethiopian village dogs were more likely to vocalise towards approaching strangers when alone in a house than when in the street. Showing that a dog's location, and if they perceive that they are guarding territory. may have an effect on their reaction in real life situations. Territorial behaviour or fear may increase aggression towards an approaching stranger depending on environment and setting and must accordingly be taken into account.

2.3.7 Handling

For this review 'handling' was defined as any behaviour test that aimed to assess how a dog coped with touch from a person. Seven papers explored hands on handling of their subjects. The tests on the subject ranged from stroking (Seksel et al 1999; Christensen et al 2007; Valsecchi et al 2011; Asher 2013), looking in the dogs mouth (Seksel et al 1999; Christensen et al 2007; Bollen and Horowitz 2008) All over body contact (Seksel et al 1999; Lucidi et al 2005; Christensen et al 2007; Bollen and Horowitz 2008; Valsecchi et al 2011; Wilsson and Sinn 2012) restraining (Asher et al 2013) or harsher manipulation (Lucidi et al 2005). The stated aims of the tests vary from finding out 'human sociability', 'response to being handled', and 'fear related behaviours', responsiveness, temperament and reaction to unknown people.

Berns et al (2015) have found that when a dog is presented with the scent of its familiar handler the caudate area in the dog's brain is activated. The caudate response represents something akin to a positive emotional response to the scent of a familiar human. This relationship between dog and handler may reduce the level of anxiety in the dog and in turn improve its response to being handled. Kerepesi et al (2015) have also found that dogs discriminate between their owner and unfamiliar people and always preferred the owner to the unfamiliar person. Therefore the results of the

handling test are naturally highly context specific as to who is performing the handling.

However in this review, in most cases an unfamiliar person handled the dog. Only Seksel et al (1999) and Wilsson and Sinn (2012) had a familiar handler present. It is debatable whether the confounding variable of the unfamiliar handler may have affected what the test was actually showing and therefore the predictive validity with respect to task performance. If the aim of the test is to predict how the dog will react to an unfamiliar person handling him or her, then the confounding variable of age and gender of the handler should be accounted for. Each test may have benefited from being repeated with different handlers of different age and gender. Seksel et al (1999) did perform a test- retest, however it is not discussed whether the evaluator conducting the second test was the same person as the prior test and none of the tests note whether the evaluator was male or female.

However if the test aims to test whether the dog can cope with just being handled, then it may have been preferable for a familiar handler to do this test, since this is the most likely type of person to handle the dog in the future.

2.3.8 Environmental sureness

A common behavioural test technique is to measure performance attribute in a novel room or novel enclosure instead of in a naturalistic setting with a familiar handler. For this to show predictive validity with respect to real life task performance then the tests must match reality. However due to the uncertainty of which type of homes many of the dogs will be placed in and with whom, this is often not possible. Nonetheless measures should be taken to ensure that test conditions are as realistic as possible. Only three studies in the review: Valsecchi et al (2011), Vas et al (2008) and Mizukoshi et al (2008) assessed their dog using a behaviour test in a real life environment for the dog.

One of the main issues with a set up situation conducted in a novel room or enclosure is that such tests have short duration compared to real life. Due to time constraints, the dogs' in novel environments may not be given sufficient time to settle in to this environment or return to normal after the previous test. The following papers did not

allow any prior habituation period for the dog: Slabbert et al 1999; Svartberg 2002; Svartberg 2005; Christensen et al 2007; Batt et al 2008, Bollen and Horowitz 2008; Vas et al 2008; Asher et al 2011; Tomkins et all 2011; Valsecchi et al 2011; Barnard et al 2012; Wilsson and Sinn 2012. Three papers did not state either way: Lucidi et al 2005, Mizukoshi et al 2008, Sforzini et al 2009. However two papers did allow a short settling in period: Sinn (2010) allowed 1-5 minutes, Seksel (2009) allowed 3 minutes. This is an important confounding variable as aversive stimuli rarely presents themselves in such quick succession in a naturally occurring environment.

Artificial conditions and succession tests can result in the subjects' emotional reserves being over stretched (Rayment et al 2015). However in real life dogs rarely go from one stressful activity straight to the next in such a short space of time. For that reason many good dogs may fail to graduate who perhaps may have performed well in a real life situation, and generally failing dogs are not given the chance to succeed later on.

2.3.9 Visual startle

Nine of the papers in this systematic review (Seksel et al 2009; Slabbert and Odendaal 1999; Svartberg 2002; Svartberg 2005; Batt et al 2008; Tomkins et al 2011; Valsecchi 2011; Barnard et al 2012; Wilson and Sinn 2012) attempted to explore how a visual startle test would affect their subjects. Some imply the startle test measures the subject's response to a novel device (Seksel et al 2009 and Barnard et al 2012), whereas Svartberg (2002, 2005) Valsecchi (2011), and Wilson and Sinn (2012) suggest that the startle test assesses avoidance, aggression, reactivity and sensitivity in their subjects. Slabbert and Odendaal (1999) go further and purport their startle test conveys how much 'spirit' and 'intellect' an animal has.

However it is precarious to attribute a given behaviour to any single emotion (Forkman et al 2007). Startle tests may measure different aspects or types of anxiety rather than just one generalised trait. Anxiety can be defined as a fear of real or imagined danger and includes depression, phobias, panic, or obsessive compulsive disorders (Costello 2007) Anxiety in a subject can build over time and may greatly influence the level of startle response that the subjects has during a test, so order effect may be important. Social isolation, a lack of habituation to the environment, order affect and what and how it is presented will all influence the degree of response.

Dogs are social animals and they often bond strongly with their own species and humans. Social isolation or an absence of a social bond with a handler may have been a highly stressful confounding variable in many of these tests. Batt et al (2008), Tomkins et al (2011), Valsecchi (2011) and Wilson and sinn (2012) all conducted their startle tests with an unfamiliar handler and in an unfamiliar environment. Although some authors do discuss a short settling in period Seksel (2009) Sinn (2010) , mostly this seems to be under 5 minutes. Apart from Seksel et al (2009) all the startle tests were also subject to order affects following on from previous tests which may have had consequences for further responses. The subject may have already formed a negative association to the environment and /or handler and already be experiencing high anxiety levels prior to the startle test, biasing the result as a consequence.

The other difficulty in analysing the results from these tests is that fear can often be displayed in different forms such as active defence (attack) active avoidance (bolting, hiding) and passive strategies (immobility) (Forkman et al 2007) . Conflict between these responses can also cause other stress related behaviours e.g. chewing on lead (Forkman et al 2007). This can lead to an array of behaviours which could be open to different interpretations. Many of the tests seem to propose that as long as the subject does not bolt or attack then it passes, despite what else it may be doing at the time. For a working dog, high degrees of anxiety, immobility or developing maladaptive (seemingly unrelated) behaviours can be extremely problematic. Therefore a pass in a startle test should not be extrapolated directly to desirable performance in real life.

Not only the nature of the aversive stimuli, but also the way and context that it is delivered can affect the startle response in an animal. Murphy et al (1981) found that when cows approached a novel object of their own free will, the ones that were most fearful of humans were the quickest to approach the object; but when the cows were forced to move toward the same novel object on leash by a human, the opposite was observed. Svartberg (2002, 2005), Tomkins et al (2011) and Barnard et al (2012) all conducted their startle tests while the dogs were on lead. Forced approach or voluntary approach all have an effect on the animal. While on lead the option for flight is removed perhaps influencing the behaviour the dog elicits.

Many of the tests where the dog was off lead use the latency to approach as the indicator of a pass, usually with more favourable scores being assigned to a quicker approach. However a non –curious/indifferent animal and a fearful animal will both show a latency to approach the same object (Forkman et al 2007). Most probably because of this, there seems to be a tendency for startle tests to use objects that are more likely to be novel to the dog. The object as well as the distance and how it is presented has an effect on the outcome of the test, with many startle test situations showing novelty but having little resemblance to everyday life. For example a dog may be startled by two men dressed as ghosts (Svartberg 2002, 2005) but the response in this situation does not indicate performance in day to day life in response to novelty. Similarly the tests that use human involvement, either when the object is a human, or has to be introduced or removed by a human, include a confounding variable. A dog that is scared by men may startle more and show greater latency to approach a man jumping out of a bush (Slabbert and Odendaal 1999) as opposed to an inanimate object rolling on the track ahead.

2.3.10 Obedience Commands

Three papers (Seksel et al 1999, Lucidi et al 2005, Valsecchi et al 2011) explored a dog's reaction to obedience commands. All three papers had an evaluator testing the dog, instead of a familiar handler. The familiarity of a dog with its handler may have affected the dog's response to the commands acting as a confounding variable. Viranyia et al (2004) found that dogs were more likely to follow a command if their owner was present and giving them visual attention. Likewise Lefebvre et al (2007) discovered that military dog's obedience was greater if they lived with their handler.

The three papers previously mentioned, also do not clarify how the command was given. Positive training methods, based on positive reinforcement have been shown to increase attentiveness toward an owner and decrease stress (Deldalle and Gaunet 2014). Therefore how the dog perceives the command is important. Rooney and Cowan (2011) discovered that dogs showed greater ability at tasks if the trainer was more playful and employed a patient approach to command giving, whereas Sumegi et al (2014) found dogs were adversely affected by their owners stress levels and performed better in working memory tasks when paired with a calm person.

Seemingly suggesting that the mood and the way a command is given has an effect on a dog's response to a command.

A further confounding variable that should be considered when testing a dog's response to commands is what type of command is given. Seksel et al (1999) use a verbal command alongside a hand signal, whereas Lucida et al 2005 and Valsecchi et al 2011 seem to only use verbal commands. Fukuzawa et al (2005) found that lip and face movements of a speaker affects a dogs perception of a command. Whereas Pongracz et al (2003) showed dogs responded better to a learnt command when just hand signals were used. Braem and Mills (2010) also interestingly found that too much verbal information or 'chat' prior to the command negatively affected performance. Thus seemingly suggesting a known hand signal given with a known one word command is the best way to avoid any confusion when testing for a dog's response to a command.

2.3.11 Noise Test

The following 13 papers (Seksel et al 1999, Slabbert and Odendaal, 1999. Svartberg,2002, 2005, Lucidi et al 2005, Christensen et al 2007, Batt et al 2008, Sforzini et al 2009, Sinn et al 2010, , Tomkins et al 2011, Valsecchi et al 2011, Wilsson and Sinn, 2012, Asher et al 2013) conducted behaviour tests to explore how their subjects behaved towards a noise. Noises ranged from; a tape recording of a thunderstorm (Seksel et al 1999), A chain with large links dragged over a sheet of corrugated metal (Svartberg, 2002, 2005), gunshot (Slabbert and Odendaal, 1999, Svartberg, K., 2002, 2005, Sinn et al 2010, Wilsson and Sinn, 2012, Valsecchi et al 2011), CD of aircraft noise (Asher et al 2013), dropped metal plate (Tomkins et al 2011, Batt et al 2013), unknown object flung to floor (Lucidi et al 2005, Sinn et al 2010) sudden barking (Sforzini et al 2009), steel buckets dropped on the floor (Wilson and Sinn, 2012), unknown sudden loud noise (Christensen et al 2007) and hairdryer (Seksel et al, 1999)

The main limitations to some noise tests, are dogs may have had previous early exposure to that particular noise already (Blackwell et al 2013). This may have either predisposed the dog to fear the noise, or habituation may have already occurred. Thus skewing any results obtained, and affecting their validity in relation to noise

sensitivity in general. For example Asher et al (2013) tests puppies with an aircraft noise, however it may be possible that this would have been heard already. The noise should ideally be one that the dog has not heard before so as to not confound the measurement.

A further confounding variable that needs to be taken into account when considering noise tests, is the location of the noise. Ideally it needs to be out of sight. Otherwise it is hard to determine if the dog is reacting to the visual or auditory stimulus. For example steel buckets dropped to the floor (Wilson and Sinn, 2012), a chain with large links dragged over corrugated metal (Svartberg, 2002, 2005), and a hairdryer (Seksel et al, 1999) may be visually frightening to some dogs, whereas the noise alone may not have affected them.

Another common problem with noise tests is identifying what part of the noise the dog is actually reacting to. For example is it the volume, the suddenness of the sound or the type of noise? All of the tests in this review have in common that the tests were sudden. Consequently perhaps it was the startle factor that the dog actually reacted to rather than the sound of the noise. If the acoustic apparatus was gradually increased in volume would the dog still react the same way?

2.3.12 Play response with a toy

Within this review, 12 papers explored dogs play response with a toy (Seksel et al 1999, Slabbert and Odendaal 1999, Svartberg, 2002, 2005, Christensen et al 2007, Batt et al 2008, Bollen and Horowitz 2008, Sforzini et al 2009, Sinn et al 2010, Valsecchi et al 2011, Wilsson and Sinn, 2012, Asher et al 2013)

McGarrity et al (2015) reminds us that tests that evaluate a dog's responsiveness to toys typically state two reasons for doing so: either to test the dogs prey/hunt/chase drive (Wilsson and Sinn, 2012, Batt et al 2008, Svartberg 2002, 2005) or to measure the dogs responsiveness/trainability (Slabbert and Odendaal 1999, Bollen and Horowitz 2008, Sforzini et al 2009, Sinn et al 2010, Valsecchi et al 2011, Wilsson and Sinn, 2012, Asher et al 2013). However Christensen et al (2007) also uses the toy test to evaluate aggression.

However as previously discussed in the ‘commands’ section, if the test is to find out responsiveness/trainability or even aggression, then the handler involved in the test may confound the measure. The test is no longer simply about its reaction to a toy. Rooney et al (2001) found that when a person ‘bows’ or ‘lunged’ at a dog during games this increased a play response from the dog independent of a toy. Rooney and Bradshaw (2010) also found after studying play sessions in fifty dog-owner partnerships, there was no evidence to suggest that dog-human games showed any link with dominance or aggression in dogs. This seems to suggest that many ‘toy tests’ are simply measuring a dogs motivation to give their attention towards a handler, and as previously discussed may be more reflective of the dogs mood state, then their ability to respond to human social cues or their ‘trainability’.

To the author’s knowledge the development of play behaviour has not been causally linked to the prey drive in domestic dogs, conversely the opposite has been found: Kittens, when given zero opportunity to play with toys, still matured into adept hunters (Hall, 1998). Suggesting animals may simply see toys as a novel object which enriches their environment (Kaulfuß and Mills 2008). However play involving toys and a play partner may provide the dog with some communicative skills between its conspecifics (Feddersen-Petersen, 2007) and provide the tester with information about the dog’s social tendencies (Feddersen-Petersen,2007). However all the dogs tested in this review were tested alone. This seems to suggest toy tests may provide more information about the dog’s reaction to novel items or their ability to entertain themselves.

2.4 General Conclusion

There appears to be very few behaviour tests in the peer reviewed scientific literature that have shown predictive validity of assessments with respect to task performance. The limitations of many are that they generalise their results to real life situations without first examining the performance of both passed and failed dogs through to real life situations. Only Valsecchi et al (2011) and Vas et al (2008) conducted a behaviour test post-graduation on both failed and passed dogs in real life settings. And only Seksel (1999) and Svartberg (2005) conducted questionnaires post-graduation on both failed and passed dogs. By excluding dogs early on, many of the assessments in

this review appear to be at best a measure of a dog's graduation potential, not whether they can actually perform the task later on.

This review has shown that many of the behaviour tests are lacking content/face validity and few are able to measure the specific character traits intended. Many suffer from order affects and subjects are more likely to be reacting to neophobia due to the unrealistic novel settings and absence of a familiar handler than demonstrate a true reflection of real life performance.

This review has included research from both working and non-working dogs, and highlighted the need for assessments with predictive validity in respect to task performance particularly in the assistance dog world, as the decisions about the future of many dogs and working partnerships are being made on the basis of inadequate tests.

Chapter 3

Analysis of the quality of behaviour tests used by Hearing Dogs For Deaf People

Contents **Page**

3.1 Introduction

The previous chapter highlighted the need for greater evaluation on both failed and passed dogs, as well as the need for greater evidence for reliability and validity within behaviour tests, especially in the field of assistance dogs. Only tests set in a natural settings, with a familiar handler were able to show any kind of content validity, with anything less being subject to the confounding variables of neophobia, order effect, blanket generalisations or at worse, stress in the subject.

The Hearing Dog tests currently in use have previously not been subject to any quality assessments, other than the in-house training with inter and intra-rater reliability measures. Therefore the aim of this study was to provide a thorough examination of the reliability of the predictive value of the T1 currently in use by Hearing Dogs For Deaf People.

The behavioural tests in use by the Assistance Dog organisation uniquely examines the behaviour of both passed and failed Hearing Dogs in naturalistic working settings, with a familiar handler present in mostly all tests. As previously mentioned, there are very few studies that both examine failed dogs past a graduation point, and are also conducted in real life scenarios. Consequently the study of Hearing Dogs behavioural tests provide a unique opportunity, but are also of vital importance for the organisation. If the test is to be relied upon in day to day life as predictive tool, the suitability and accuracy of the 8 month test must first be investigated.

The hypothesis of this study was that Performance in the 18 month Hearing dog test (T2) can be predicted from the dog's performance in their 8 month (T1) Hearing Dog test.

The aim of this study was to investigate the relationship between the results from both T1 and T2.

The objective was to extract data from 62 dog's behaviour tests on both their T1 and T2. This was done in order to investigate each behavioural factor separately; firstly examining any correlations or differences between the 8 and 18 month data sets, followed by an investigation of the sensitivity and specificity of each behaviour test. Where appropriate the convergence level of the behavioural tests were measured

against a CBARQ questionnaire. In light of the results a thorough critique of each test is provided.

3.2 Methods

This study was approved by the relevant University ethics committee. Hearing Dogs for Deaf People gave written informed consent (see appendix 7), as did James Serpell for the CBARQ questionnaire (Hsu and Serpell 2003) (see appendix 6).

3.2.1 Subjects

A total of 62 dogs participated in this study. All the dogs that took part in the two Hearing Dog tests had been in the Hearing Dog program since they were 8 weeks old. The sample consisted of 17 Labradors, 2 Labrador cross golden retrievers, 6 miniature poodles, 14 show cockers, 9 working cocker spaniels, 9 working cocker spaniel cross poodles and 5 working cocker spaniels cross show cocker spaniels. 38 males and 24 females were in the final sample group. Dogs were chosen for the sample if they were due to have an 8 month (T1) and 18 month (T2) assessment within the time span of this thesis. Any dogs that entered the breeding program before having their T2 or had health problems were removed from the original sample group.

4 males and 4 females had been neutered prior to the 8 month test (54 dogs had not been neutered prior to the 8 month test). The whole sample group (62 dogs) had been neutered prior to the second test. For the first test there was a median age of 8 months and a modal age of 8 months old. For the second test there was a median age of 18 months old and a modal age of 18 months old. Out of 62 dogs 23 dogs were transferred at approximately 14 months old to the Northern Hearing Dog site to begin their advanced training. 27 were transferred from their volunteer socialiser at 14 months old to be trained at the Southern Hearing Dog Site. 12 dogs carried out advanced training from 14 months old from their familiar volunteer socialiser's home.

From Monday to Friday all dogs in the sample, irrespective of where they were being trained, completed 2 hours of training a day by a trained member of staff. The training consisted of approximately half an hour of sound work training, half an hour of town training and 1 hour of off lead park training. All dogs that were being trained

on a training site were collected by a volunteer socialiser between 5.30 to 6.30 pm every weekday evenings, and then dropped back again the following morning between 6.30 to 8.30 am. All 62 sample dogs had the weekend off resting and going for casual walks with their volunteer. (Each volunteer consistently had the same dog for the whole of that dog's advanced training period). Occasionally if a weekend volunteer was sick or away the dog would spend its' weekend with a different volunteer socialiser, however this never spanned over a period of time longer than 2 weekends and did not occur when the weeks that the dogs were tested.

3.2.2 The Hearing Dog Test

Dogs were rated on a scale between 1-5, with 1 representing a poor performance and 5 representing a high performance for the 13 behaviour factors of 'social behaviour (adults), social behaviour (children), social behaviour (dogs), environmental behaviour, recovery rate, adaptability, vocal reactivity, motivation, trainability, frustration, chase hunt and distractibility. The behaviour tests began any time from 10.30 am and followed normal Hearing Dog day to day procedures with one of the four Hearing Dog quality assurance evaluators scoring the test (See introduction, pages 8-20 for full Hearing dog test procedures, evaluator information and scoring). 20 of the dogs had the same QAE for the first and second test. Forty two dogs had different QAE's for each test.

3.2.3 The questionnaire

Socialisers and trainers were asked to complete a CBARQ (Canine Behavioural Assessment and Research Questionnaire) questionnaire (Hsu and Serpell, 2003) that required them to answer 100 behavioural questions on the Hearing Dog in their care. When the dogs were 7 months old the Socialisers were requested by post with a covering letter (see appendix 4) to complete the first questionnaire (CBARQ-1) as soon as possible. This was chased up by phone call by the researcher when the dog was 8 months old, and then again at 9 months old to any people who had not yet returned it. The trainers were then asked to complete the same questionnaire when the dog had been in advanced training for two months (approximately 16 months old) (CBARQ-2). This was chased up by an email one month later when the dogs were

approximately 17 months old and then again a month after that when they were approximately 18 months old. This was done by either a phone call if they worked at the Northern site, or where possible face to face if they worked at the southern site. The overall return rate on the questionnaires was poor, with many that were returned missing names or missing fields. Only 12 useful pairs of questionnaires could be used for this study.

The following CBARQ factors: 'Dog-directed aggression', 'Dog rivalry', 'Stranger-directed fear', 'Non-social fear', 'Attachment and attention-seeking', 'Trainability' and 'Chasing' were compared with the corresponding Hearing Dog behavioural factors for: 'social behaviour (dogs)', 'social behaviour (adults)', 'environmental behaviour', 'adaptability', 'trainability' and 'chase'. The trainers that filled in the later questionnaire were blind to the information that was supplied by the socialiser in the preceding questionnaire, although they were privy to past knowledge of the dog's behaviour by other sources of information. Scores were rated on a scale between 0 and 4 with 0 being a good performance 4 representing a poor performance. A copy of the CBARQ questionnaire and how the scores were further calculated are presented in the Appendix (See appendix 3 and 4).

3.2.4 Statistical Analysis

Statistical analysis was performed using IBM SPSS version 22 software. For the evaluation of the predictive value between T1 and T2, two initial sets of tests were performed; firstly to determine the strength of the relationship between T1 and T2 a Spearman's correlation was performed. Secondly to answer the question; Is there a significance difference in scores of the T1 and T2, a Wilcoxon matched pairs test was applied to the two data sets (T1 and T2). For a further assessment of the Hearing dog behavioural test, sensitivity and specificity controls was applied. Sensitivity is the ability of a test to correctly identify a true positive rate, whereas specificity is the ability of a test to correctly identify a true negative rate. Positive and negative likelihood ratios were also applied to T1 and T2s test data. Likelihood ratios were used in this study to provide information on the probability of a dog passing given the number of passes and failures in the population. Due to the relatively small size and the exploratory nature of the study, and despite multiple testing, a probability threshold of 0.05 was used. Accordingly the result must be treated with caution.

To estimate the convergent validity between the CBARQ questionnaires and Hearing Dog test scores a Spearman's Rho correlation was applied between T1 compared with CBARQ-1 and T2 test data and CBARQ-2.

To ascertain the predictive value of the opinions of the socialisers a Spearman's Rho correlation was calculated to the corresponding CBARQ-1 compared with the CBARQ-2 score. A Wilcoxon matched pairs was also used to determine the significant different between the CBARQ-1 and the CBARQ-2 scores.

3.3 Results

The sample group consisted of 62 dogs. All 62 took part in both T1 and T2 (see table 4). The median age of dogs taking part in the first test was 8 months old. The median age of dogs taking part in the second test was 18 months old.

3.3.1 Comparisons between Hearing dog behaviour tests T1 and T2

In order to test for predictive validity of the second Hearing Dog behaviour test from the first Hearing Dog Behaviour test, comparisons were made between the two different data sets.

3.3.1.1 Correlations of T1 and T2 (see Table 6)

A Spearman's Rho correlation was conducted on the data from the first Hearing Dog test compared with the second hearing dog test. Table 6 shows a significant but moderate positive correlation between 'children' (r^S =.419, p=.000), 'adaptability' (r^S =.404, p=.001), 'vocal reaction' (r^S =.400, p=.001), 'chase' (r^S =.434, p=.000), and 'hunt' (r^S =.463, p=.000). Meaning there is a moderate linear relationship between the two Hearing Dog behaviour test's for these behavioural factors. All the relationships were positive i.e: as test one scores increased, test two scores also increased.

Table 6 shows that a significant but weak positive correlation was found between test one and two for: 'adults' (r^S =.244, p=. 028) , ' dogs' (r^S =.368, p=.002) , 'environmental behaviour' (r^S =.298, p=.009), 'motivation' (r^S =.383, p=.001), 'trainability' (r^S =.361, p=.002) and 'frustration' (r^S =.327, p=.005) Meaning that although statistically significant, the strength of these linear relationships were weaker. Therefore due to the relatively small sample size caution should be taken before concluding there is a reliable association between these variables.

The Spearman's correlation coefficients between the Hearing Dog test one and two in the area of 'Distraction' *(r^S = .158, p=.110)* and' recovery rate'(r^S = .206, p=.054)

were not significant at the .05 probability level, indicating poor agreement between the tests.

Table 6: Table to show Spearman's Rho test results, medians and probability values for 8 and 18 month Hearing Dog Behavioural tests

Behaviour test	N		Minimum		Maximum		Median		Spearman's Rho	Probability
	8 month	18 month	8 month	18 month	8 month	18 month	8 month	18 month		
Social behaviour - adults	62	62	2	2	5	5	4	4	0.244	0.028
Social behaviour - children	62	62	2	2	5	5	4	4	0.419	0.00
Social behaviour - dogs	62	62	2	2	5	5	3.5	3	0.368	0.002
Environ-mental behaviour	62	62	2	2	5	5	4	4	0.298	0.009
Recovery rate	62	62	2	2	4	4	4	3	0.206	0.054
Adaptability	62	62	2	2	5	4	4	3	0.404	0.001
Vocal reactivity	62	62	2	2	5	5	3	3	0.400	0.001
Frustration	62	62	2	2	5	4	3	3	0.327	0.005
Motivation	62	62	2	3	5	5	4	4	0.383	0.001
Trainability	62	61	3	1	5	5	4	4	0.361	0.002
Chase	62	62	2	3	5	5	4	4	0.434	0.00
Hunt	62	62	3	3	5	5	4	4	0.463	0.00
Distraction	62	62	2	2	4	4	3	3	0.158	0.11

Fig 3 Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'social behaviour –adults'. (Size of bubble indicates number of responses with each value)

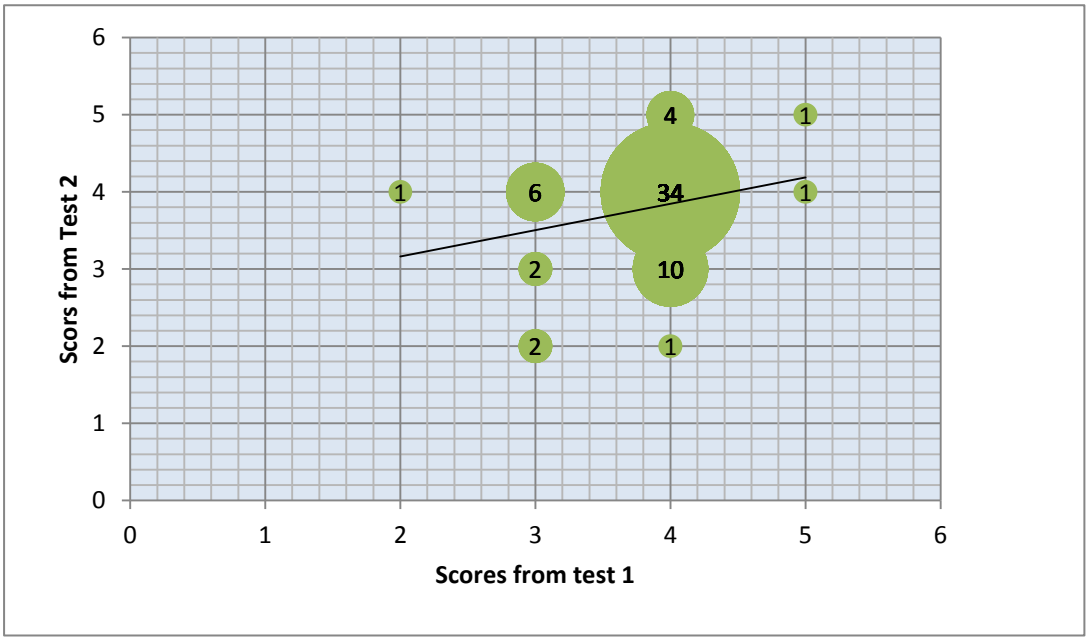


Fig 4 Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'social behaviour-children'(Size of bubble indicates number of responses with each value)

Fig 5 Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'social behaviour-dog' (Size of bubble indicates number of responses with each value).

Fig 6 Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'environmental confidence' (Size of bubble indicates number of responses with each value)

Fig 7 Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'Recovery rate '(Size of bubble indicates number of responses with each value)

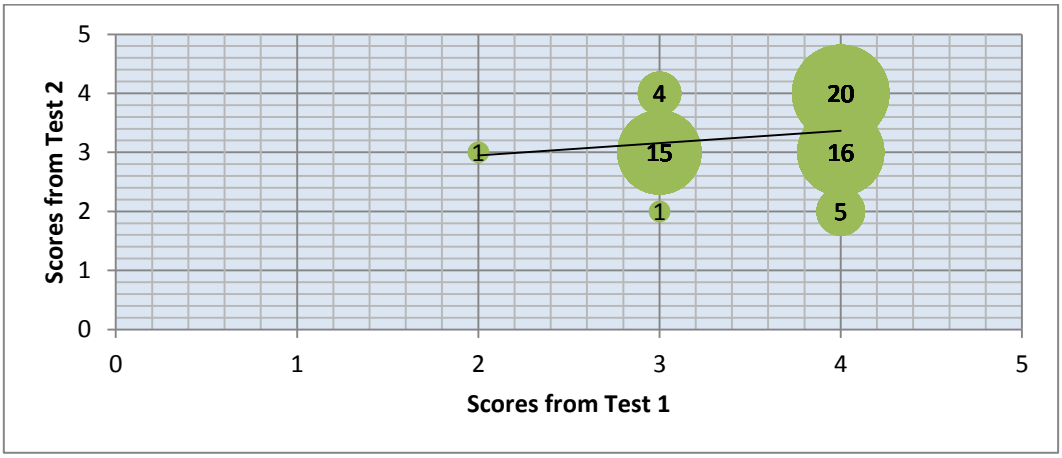


Fig 8 *Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'Adaptability' (Size of bubble indicates number of responses with each value)*

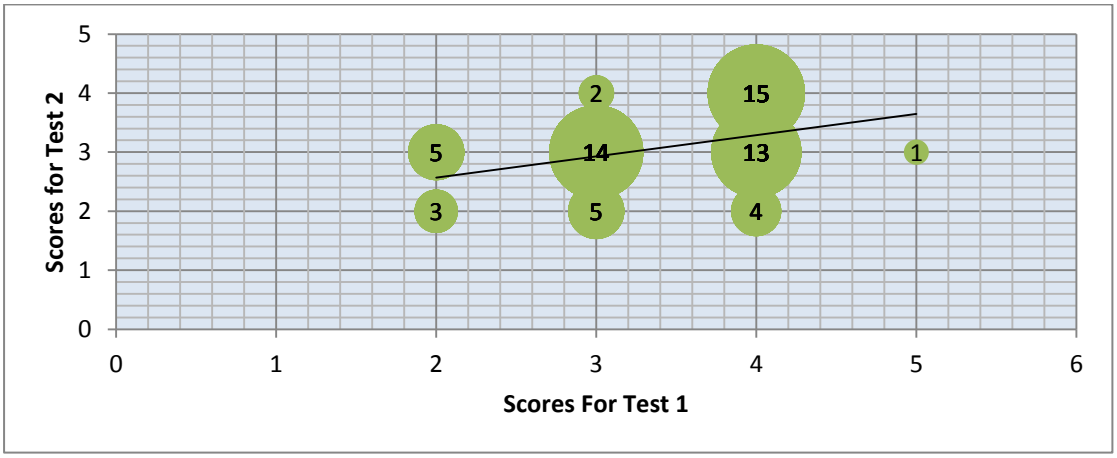


Fig 9 *Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'vocal reactivity' (Size of bubble indicates number of responses with each value)*

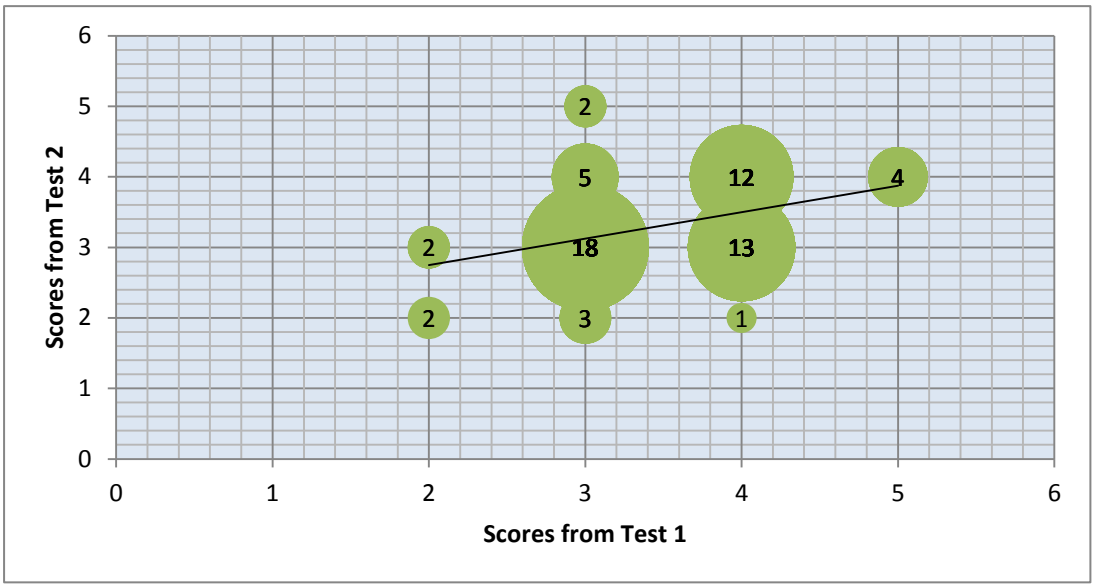


Fig 10 *Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'frustration' (Size of bubble indicates number of responses with each value)*

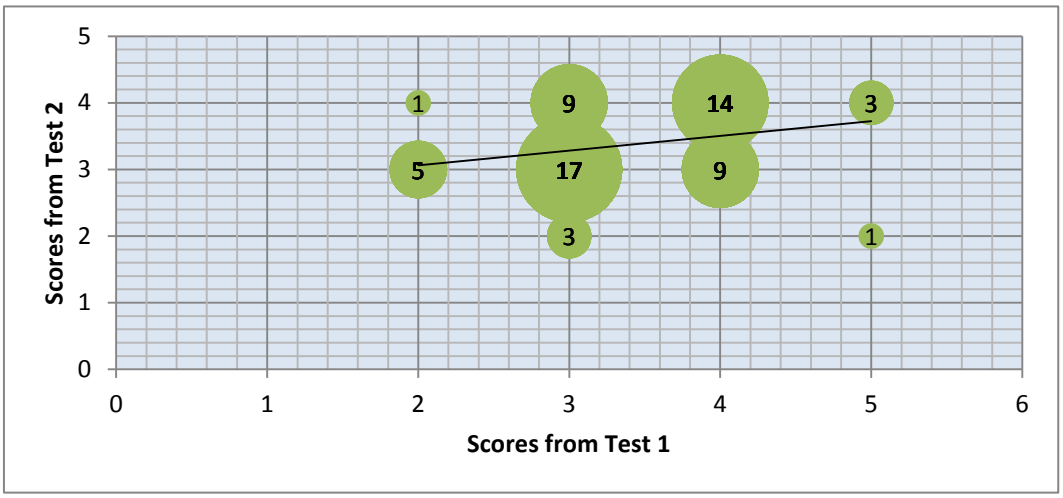


Fig 11 *Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'motivation '(Size of bubble indicates number of responses with each value)*

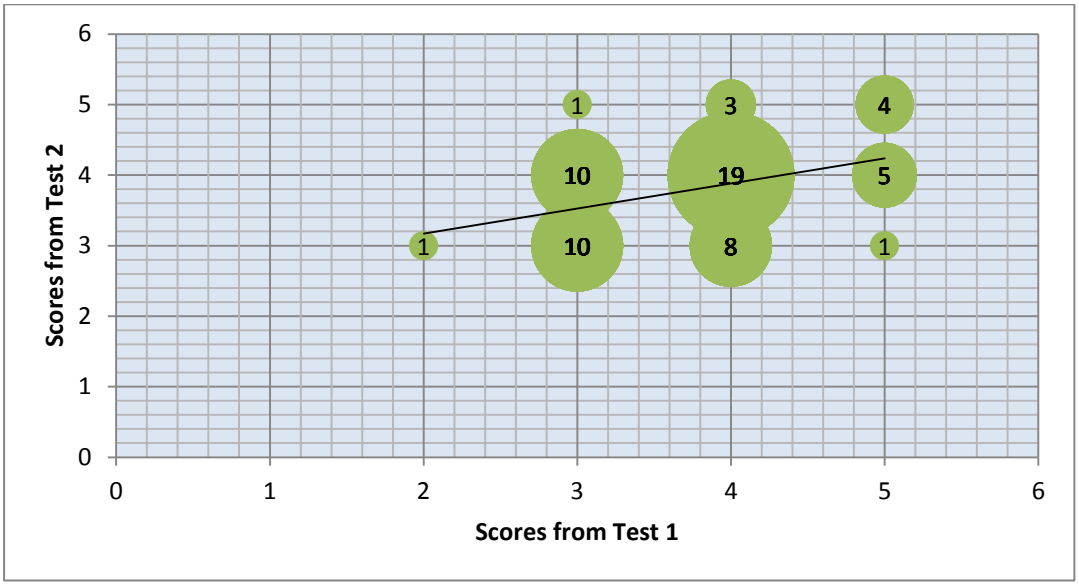


Fig 12 *Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'Trainability' (Size of bubble indicates number of responses with each value)*

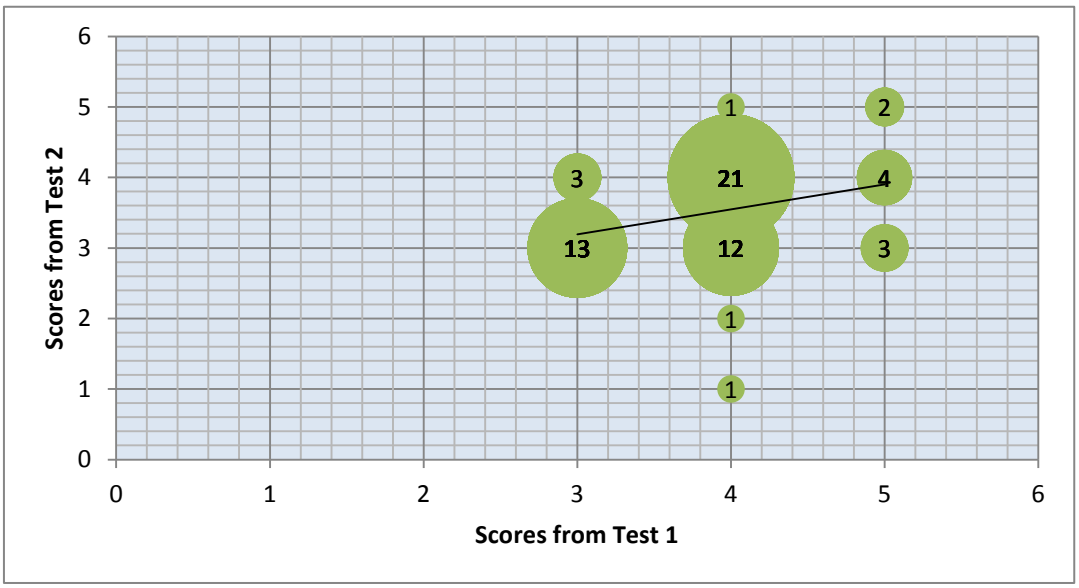


Fig 13 *Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'Chase'(Size of bubble indicates number of responses with each value)*

Fig 14 *Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'Hunt' (Size of bubble indicates number of responses with each value)*

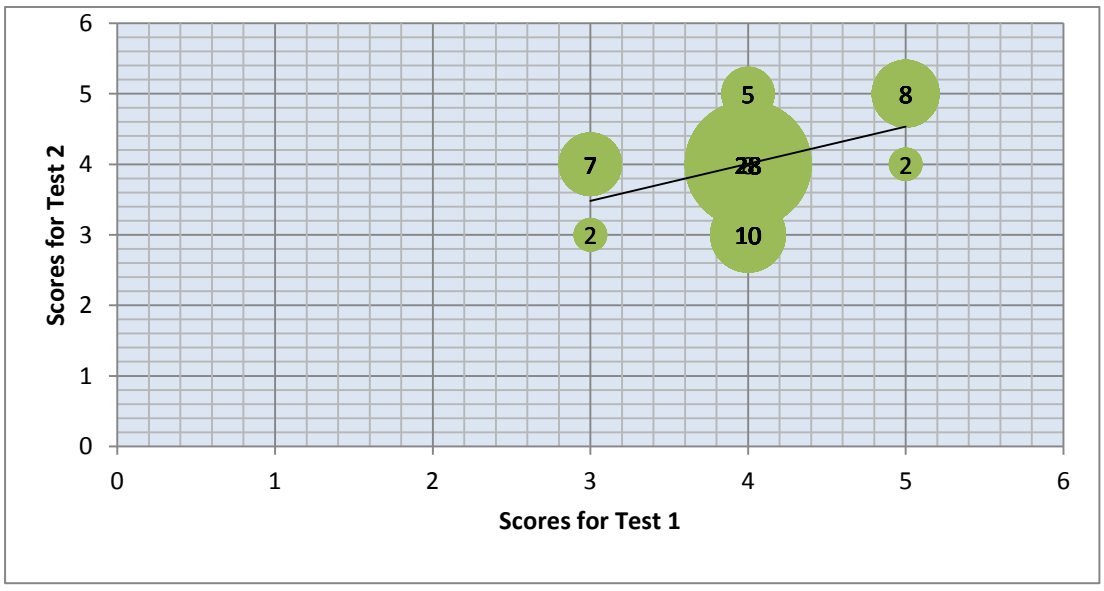
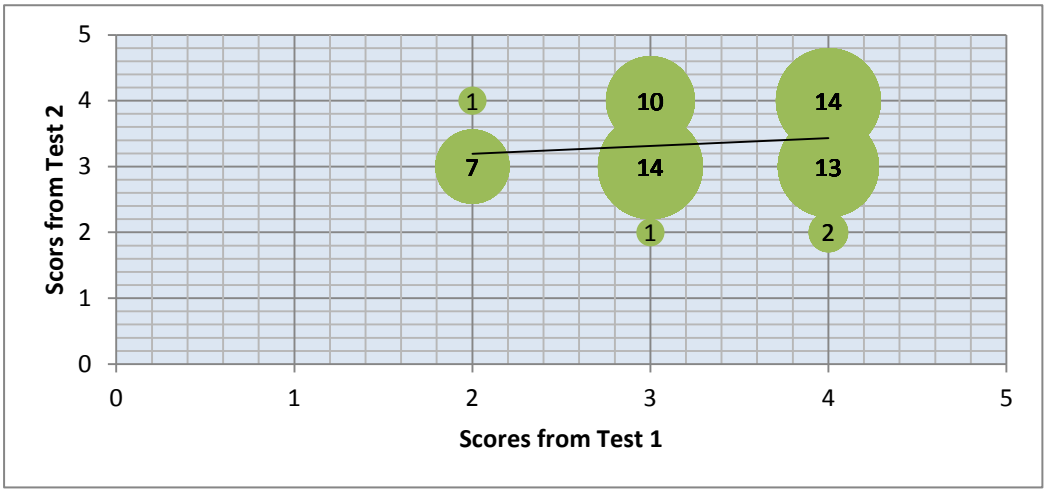


Fig 15 *Bubble chart to show frequency of 8 and 18 month scores for the Hearing Dog test for 'Distractibility'(Size of bubble indicates number of responses with each value)*



3.3.2 Differences **(see Table 7)**

The Spearmans Rho is able to show the linear relationship between two variables, however it is not able to quantify the relative size of the values between the two tests, only that they tend to generally move in the same direction. Difference between pairs of scores obtained in test one or two for each behaviour factor was further analysed using a Wilcoxon matched pairs test.

Differences between the medians of test one and test two were analysed (see Table 7) using a p-value < .05. The Wilcoxon test demonstrated that the median scores obtained for the behavioural factor of 'recovery rate' *(z= -3.386,p =.001, r=11.14)* , 'adaptability' *(z= -3.136, p =.002, r= 11.14)* and 'trainability' *(Z= -3.459, p=.000, r=11.14)* significantly declined between test one and two (see Table 4)

At first glance this may seem to contradict the Spearmans Rho findings. However paired tests may decline in their median, while the Spearmans Rho of the population remains positively correlated. This is because the Wilcoxon is looking at the pseudo-median (not the true population median) of the pairwise averages. It is often referred to as the median but the two can be very different when using a non symmetric population such as this research does. For example if we firstly look at 'Adaptability' (see fig 8) the trend of a positive correlation is influenced by 2 points on the x-axis

where the scores for the second test are higher; 5 dogs obtained a score of 2 on their first test, but then a score of 3 in the second test. Similarly 2 dogs scored a 3 on their first test but then scored a 4 on their second test. These scores along with the data of 14 dogs scoring 3 on both their first and second test and 15 scoring 4 on both their first and second test would have yielded the Spearmans Rho results.

However there are four points on the axes where the scores for the second test are lower than the first; One dog scored 5 on their first test but then dropped 2 whole units and only scored 3 on the second test. 5 dogs scored 3 on their first test, but only a 2 on their second. 4 dogs scored 4 on their first test, but only 2 on their second test, and 13 dogs scored 4 on the first test but then only 3 on their second test. While there is a larger number of subjects scoring the same or higher in their second test, there are still four points on the axes and 22 dogs where their tests declined. Because there are more data points lower for those scoring 4 in the first test (and some of them are dropping by 2 units) this means the difference overall is largely negative. It must also be taken into account that for the largest majority of the data the dogs score the same in the two tests, so this may also partly be the reason for the differences between the statistical outcomes.

The behavioural factor of 'Trainability' has only 61 dogs ranked scores, as one dog is missing their T2 score. This of course may have skewed the results slightly and should be taken into consideration. Overall the correlation trend was positive, however the significant negative Wilcoxon was obtained due to 21 dogs having ranks that decreased from T1 to T2. (see fig 12) For example one dog dropped his scores by three scores , (this dog went from a score of 4 to a score of 1) , 4 dogs dropped their scores by two (1 dog went from a score of 4 to a 2, and 3 dogs went from a 5 to a 3) and 16 dogs dropped their scores by one 1 (12 scored a 4 in T1 but only a 3 in T2; 4 scored a 5 in T1 and only a 4 in T2). In contrast only 4 dogs ranks increased from T1 to T2 (3 dogs went from a score of 3 on their T1 to a 4 on their T2; 1 dog went from a 4 in T1 to a 5 in T2). The rest of the ranked scores were tied and came from 36 dogs that scored both the same on their T1 as they did on their T2.

When looking at the scores for 'Recovery Rate' (see fig 7) there is not a significant Spearmans Rho correlation, however there is a significant difference. There are 22 data points where the dogs score lower on their second test than they did on their first. The majority of these come from 16 dogs who scored 4 on their first test but only a 3

on their second. This is perhaps explaining why a significant negative difference has been found between these ranked pairs of scores. However the large number of dogs that scored the same in their first test as they did in their second (35) may skew the overall results slightly.

There was no significant difference between the medians of test one and two's behavioural factors for: 'Social adults, 'social children', 'dogs', 'environmental', 'vocals', 'frustration', 'motivation', 'chase', 'hunt' and 'distraction'(see Table 7)

Table 7: Table to show Wilcoxon test results, medians and probability values for 8 and 18 month Hearing Dog Behavioural tests

Behaviour test	N		Minimum		Maximum		Median		Z Score	Probability
	8 month	18 month	8 month	18 month	8 month	18 month	8 month	18 month		
Social behaviour - adults	62	62	2	2	5	5	4.00	4.00	-0.536	.693
Social behaviour - children	62	62	2	2	5	5	4.00	4.00	-1.895	.072
Social behaviour - dogs	62	62	2	2	5	5	3.50	3.00	-1.768	.110
Environmental behaviour	62	62	2	2	5	5	4.00	4.00	-1.024	.382
Recovery rate	62	62	2	2	4	4	4.00	3.00	-3.386	.001
Adaptability	62	62	2	2	5	4	4.00	3.00	-3.136	.002
Vocal reactivity	62	62	2	2	5	5	3.00	3.00	-1.731	.108
Frustration	62	62	2	2	5	4	3.00	3.00	-0.175	.879
Motivation	62	62	2	3	5	5	4.00	4.00	-0.168	.951
Trainability	62	61	3	1	5	5	4.00	4.00	-3.459	.000
Chase	62	62	2	3	5	5	4.00	4.00	-1.512	.185
Hunt	62	62	3	3	5	5	4.00	4.00	.000	1.000
Distraction	62	62	2	2	4	4	3.00	3.00	-0.142	.984

3.3.3 Overall Comments

A review of the two previous statistical tests using data from test one and two, seemingly suggests that the behavioural factors of ‘children’, ‘vocal reaction’, ‘chase’ and ‘hunt’ demonstrate a moderate positive correlation with an absence of any significant differences between test one and two’s paired scores.

Whereas the behavioural factors of ‘adults’, ‘ dogs’, ‘environmental behaviour’ ‘motivation’ and ‘frustration’ show a significant but weak correlation with an absence of any significant differences between paired scores.

Adaptability has a significant moderate positive correlation and is showing a significant decline between test T1 and T2s paired scores. Whereas, ‘trainability’ has a significant but weak correlation with significant decline between its paired scores.

Recovery rate is not demonstrating a significant correlation, and it is showing significant decline for the differences between the two data sets.

Distractibility is showing neither a significant correlation, nor any significant differences.

3.3.4 Sensitivity and specificity of Hearing Dog behaviour tests

(See Table 8)

The sensitivity and specificity of the first Hearing Dog test as a predictor of results from the comparable second behaviour tests were calculated for the putative areas of social behaviour with ‘adults’, ‘children’ and ‘dogs’, ‘environmental behaviour’, ‘Recovery rate’, ‘adaptability’, ‘vocal reactivity’, ‘frustration’, ‘motivation’, ‘trainability’, ‘chase’, ‘hunt’ and ‘distraction’. Looking at the sensitivity and specificity measures (see Table 8), with most testing 0 % for specificity and only two: adaptability scoring 25% and vocal reactivity scoring 33% for specificity this demonstrates an overall high sensitivity and low specificity of the Hearing Dog behaviour tests. (For full description of each test see appendix 5)

The negative likelihood ratios were mostly incalculable due to the 0% specificity, however ‘adaptability’ had a negative likelihood of 0.4 and vocal reactivity had a

negative likelihood ratio of '0.12'. The positive likelihood ratios were: social behaviour with 'adults'(0.99), 'children' (0.99) and 'dogs'(0.98), 'environmental behaviour'(0.97), 'Recovery rate' (0.98), 'adaptability'(1.2), 'vocal reactivity'(1.43), 'frustration'(0.9), 'motivation'(0.98), 'trainability'(1), 'chase'(0.49), 'hunt' (1)and 'distraction' (6.14).

Table 8: Table to show Sensitivity, Specificity and likelihood measures for Hearing Dog Behaviour tests.

Hearing Dog Behaviour Test	Sensitivity	Specificity	Likelihood ratio- Positive	Likelihood ratio- Negative
Social behaviour-adults	99%	0%	0.99	not calculable
Social behaviour-children	99%	0%	0.99	not calculable
Social behaviour-dogs	98%	0%	0.98	not calculable
Environmental-behaviour	97%	0%	0.97	not calculable
Recovery rate	98%	0%	0.98	not calculable
Adaptability	90%	25%	1.2	0.4
Vocal reactivity	96%	33%	1.43	0.12
Motivation	98%	0%	0.98	not calculable
Trainability	100%	0%	1	not calculable
Frustration	90%	0%	0.9	not calculable
Chase	98%	0%	0.49	not calculable
Hunt	100%	0%	1	not calculable
Distractibility	86%	0%	6.14	not calculable

3.3.5 Convergence

To test for convergence, 12 pairs of results from the first and second Hearing Dog behaviour test were compared with the concomitant sections in 12 pairs of returned CBARQ questionnaires for the same dogs.

3.3.5.1. 8 T1 compared with CBARQ-1 questionnaire (see Table 9)

There was a significant correlation coefficient between CBARQ-1 'stranger directed fear' compared with the T1 'adult' (r_s, = -.516, p=.043). The relationship is negative: as CBARQ-1 scores increased, T1 decreased. However as a high number in the

CBARQ means the dog did badly and a low number in the Hearing Dog test means the dog did badly the hypothesis is supported for this one measure i.e: the Hearing Dog T1 for 'social behaviour-adults' appears to measure something similar to stranger directed aggression in the CBARQ-1.

As shown in Table 9, there was no other significant correlations found between any HD test and any corresponding CBARQ component.

3.3.5.2 T2 compared with CBARQ-2 (see Table 9)

The sig value for the Spearman's correlation coefficient between 'stranger directed aggression' CBARQ-2 questionnaire compared with the T2 'adult' Hearing Dog assessment is $r^s = -.568$, $p = .034$; therefore it can be concluded there is a significant relationship between the 'stranger directed aggression' CBARQ-2 questionnaire and the T2 for 'social behaviour-adult'.

The relationship is negative: as CBARQ-2 scores increased, T2 decreased. However as a high number in the CBARQ means the dog did badly and a low number in the hearing dog test means the dog did badly the hypothesis is supported: As the dogs scores improved on the C BARQ-2 test so did they improve on the T2 behaviour test.

There were no further significant correlations found between the T2 compared with concomitant CBARQ-2 scores, as can be seen in Table 9 .

3.3.5.3. T1 compared with CBARQ-2 (see Table 9)

Using a Spearman correlation test there were no significant correlations between any of the scores obtained from T1 and CBARQ-2 questionnaires.

3.3.5.4. CBARQ-1 compared with T2 (see Table 9)

A Spearman's Rho Correlation is significant at the p= 0.05 level for the CBARQ-1 Non-social fear compared with the T2 Environmental test (r^s =-499, p=.049); the CBARQ-1 trainability question compared with the T2 test (r^s =.513, p=.044), and the attachment and attention seeking CBARQ-1 questionnaire score and T2 adaptability score *(r^s =-716, p=.004)*. The correlations for trainability scores is positive, meaning as the CBARQ-1 scores increase so do the T2 scores. However as a high number in the CBARQ means the dogs scored poorly but a high number in the Hearing Dog test means the dog scored positively, these results actually show that if dogs scored

positively on the CBARQ-1 questionnaire in the areas of trainability they actually scored negatively on the corresponding T2 test for trainability. However in the CBARQ-1 ‘non-social fear’ compared with T2 environmental test and the CBARQ-1 ‘attachment and attention’ seeking compared with the T2 ‘adaptability’ test the relationship is negative. This means as CBARQ-1 scores increased, T2 scores decreased. However as a high number in the CBARQ means the dog did badly and a high number in the hearing dog test means the dog did well for the areas of environmental confidence and adaptability the hypothesis is supported: As the dogs scores improved on the CBARQ-1 test so did they improve on the T2.

The tests were non significant for all other comparisons between the CBARQ-1 and the Hearing Dog T2 (see Table 9)

Table 9: Table to show Spearman's Rho test results, medians and probability values for the 8 and 18 month Hearing Dog behavioural tests compared with concomitant 8 and 18 months CBARQ scores.

Behaviour test	N		N		Median		Median		(Spearmans Rho) and Probability			
	T1	CBARQ-1	T2	CBARQ-2	T1	CBARQ-1	T2	CBARQ-2	T1 vs CBARQ-1	T2 vs CBARQ-1	T1 vs CBARQ-2	T2 vs CBARQ-2
CBARQ 'Stranger directed aggression'/ HD 'Adult' test	12	12	12	11	4.0	.050	3.5	.300	(-0.449) 0.071	(-0.485) .055	(-0.407) 0.107	(-0.568) 0.034
CBARQ 'Stranger directed fear'/ HD 'Adult' test	12	12	12	12	4.0	0.00	3.5	.375	(-0.516) 0.043	(-.477) .059	(-0.477) 0.073	(-0.248) 0.219
CBARQ 'Dog directed aggression'/ HD 'Dog test'	12	12	12	12	3.5	0.00	3.0	0.00	(-0.302) 0.170	(-.095) .384	(-0.109) 0.368	(-0.498) 0.053
CBARQ 'Dog rivalry'/HD 'Dog' test	12	4	12	7	3.5	0.00	3.0	.500	(-0.333) 0.333	(-0.577) .211	(-0.294) 0.261	(-0.521) 0.116
CBARQ 'Dog directed fear/ HD 'Dog' test	12	12	12	10	3.5	.250	3.0	.625	(-0.103) 0.374	(-0.27) .191	(-.497) .072	(-0.071) 0.847
CBARQ Non-social fear'/HD 'Environmental' test	12	12	12	11	3.5	.330	3.5	.670	(-0.123) 0.352	(-499) 0.049	(-0.493) 0.062	(0.000) 0.500
CBARQ 'attachment and attention seeking'/HD 'Adaptability' test	12	12	12	12	4.0	1.67	3.0	2.00	(-0.488) 0.054	(-716) 0.004	(-0.04) 0.451	(-0.388) 0.106
CBARQ 'Trainability test'/ HD 'Trainability' test	12	12	12	12	4.0	2.82	3.0	2.13	(-0.516) 0.439	(-0.513) 0.044	(346) 0.135	(-0.292) .178
CBARQ 'Chase'/HD 'Chase' test	12	8	12	10	4.0	1.63	4.0	1.63	(-0.321) 0.219	(-0.22) .298	(-0.231) 0.261	(-0.14) .346

3.3.5 Socialisers opinions compared with Trainers opinions

The CBARQ was utilised to further test if opinions between socialisers and trainers correlated or differed for the various behavioural factors.

3.3.5.1. CBARQ-1 compared with CBARQ-2 (see Table 10)

The results for the 1- tailed Spearman's correlation coefficient in CBARQ-1 and CBARQ-2 relating to 'stranger directed aggression' (r^s =.800, p=.002) and 'dog directed fear' (r^s =.601, p=.033) were less than 0.05; therefore it can be concluded that in the areas of stranger directed aggression and dog directed fear there is a significant correlation between the CBARQ-1 and CBARQ-2 scores. The relationship for stranger directed aggression and dog directed fear is positive: CBARQ-1 scores increased, so did the CBARQ-2 scores.

Using a Wilcoxon matched pairs test (see Table 8), 'Stranger-directed aggression' *(Z= -2.388,p=.016, r= 4.9),* 'Stranger-directed fear' *(Z= -2.384, p= .016, r= 4.9)* , 'non-social fear' *(Z= -2.315, p=.019, r=4.9),* and 'Trainability' *(Z=-2.752,p=.004, r=4.9)* all show significant difference between the CBARQ 1 and CBARQ-2.

It can be concluded (based on the fact that negative ranks were used) there was a significant decrease in scores in the CBARQ-2 compared with the CBARQ-1 in 'stranger directed aggression', 'stranger directed fear' and 'non-social fear'. This means the reported cases of 'Stranger-directed aggression', 'Stranger-directed fear' and 'non-social fear' had all decreased by the CBARQ-2 questionnaire. However in the case of' trainability' (based on the fact that positive ranks were used) there was a significant increase in scores in the CBARQ-2 compared with the CBARQ-1. Therefore dogs were marked more negatively for their trainability and their energy levels in their CBARQ-2 questionnaire compared with their CBARQ-1 questionnaire

As can be viewed in Table 10, there were no further significant results between CBARQ-1 and CBARQ-2.

Table 10: Table to show Wilcoxon test results, medians and probability values for 8 and 18 month CBARQ questionnaire.

CBARQ Section	N		Minimum		Maximum		Median		Z Score	Probability
	8 month	18 month	8 month	18 month	8 month	18 month	8 month	18 month		
Stranger directed aggression	12	11	0.00	0.00	.80	1.70	.0500	.3000	-2.388	.016
Owner-directed aggression	12	11	0.00	0.00	.50	.63	0.0000	0.0000	-.816	.750
Dog-directed aggression	12	12	0.00	0.00	.50	1.00	0.0000	0.0000	-1.511	.250
Dog rivalry	4	7	0.00	0.00	.25	1.75	0.0000	.5000	-.447	1.000
Stranger-directed fear	12	12	0.00	0.00	.25	1.50	0.0000	.5000	-2.384	.016
Non-social fear	12	11	0.00	.17	1.17	1.70	.3300	.6700	-2.315	.019
Dog-directed fear	12	10	0.00	0.00	1.00	2.25	.2500	.6250	-1.825	.086
Attachment and attention-seeking	12	12	1.00	1.17	1.17	3.17	1.6700	2.0000	-1.219	.240
Trainability	12	12	1.75	1.38	1.38	3.00	2.8150	2.1300	-2.752	.004
Chasing	8	10	.50	.75	.75	3.00	1.6250	1.6250	-.425	.813

3.3.6 Case Studies

Due to the scope of this thesis it is not possible to complete a case study on all 62 dogs. However it may be beneficial to look at the case studies of those that failed either T1 or T2 within the factors with the stronger correlations and no significant differences: ‘children’, ‘vocal reaction’ and ‘chase’. Interestingly ‘hunt’ is also one of the factors that had a moderate correlation, but no differences. However no dogs failed at either T1 or T2.

3.3.6.1 Children

For the behaviour test of 'Children', two dogs failed on T1 (obtained a scores of 2 or lower). One was a cocker cross poodle and one was a Labrador. Both dogs had the same QAE on T1 and were tested at the Southern centre. Both dogs were unneutered on T1 and had not had their first season yet. Both dogs received extra training and support with children throughout their socialising period, with this increasing after their first test. All additional support was done in their own homes or surrounding area up to the age of 14 months. Two months before being transferred to a training centre both dogs were neutered. After this, the Labrador was transferred to the Northern centre, and the Cocker spaniel cross Poodle was transferred to the Southern training centre. Continued support and training was given at these centres.

On T2 both dogs were tested by a different QAE from each other and from the previous test. By T2 the Labrador that was re tested at the Northern centre had improved around children and increased her score from a 2 to a 3. This dog is now out working as a full Assistance Hearing Dog. The Cocker spaniel did not improve, and continued to have extreme anxiety around children, despite extensive positive behaviour modification. After failing the second test she was re-homed to an experienced Hearing Dog volunteer who has no contact with children, but who is continuing behaviour modification in this area with the help of an experienced behaviourist.

Two dogs passed T1 with children (obtained a score of 3 or higher) but then failed in T2. Both dogs went from a score of 4 on their first test to a score of 2 on their second test. Both dogs had a different QAE for T1 compared with T2. There was a different QAE for both of these dogs at T1 and T2. For the purposes of this thesis we will call dog one Dusty and dog two Terry. Dusty is a Cocker spaniel from working lines. Dusty's problem began to develop two months after T1 when she unfortunately had a negative incident with an unfamiliar four year old child in her home (which was not the dog's fault) Since then anxiety around unfamiliar children increased in frequency, to a point where she would growl, show teeth, and show strong avoidance around any small child in the home, in town, or in the park (Her behaviour never escalated past this point). Dusty did receive positive behavioural support immediately after the

incident, but unfortunately never fully became trust worthy around children again. After failing her second test she received a change of career within Hearing Dogs and became a 'confidence and companion' dog. According to organisational policies Dusty should have had a change of career and become an SSD (sound Support Dog) as she had passed her sound work test, however at the time there were no child free environments on the waiting list for this career role. She was consequently re-homed with a client that has a minor hearing loss, and lives on her own with no family and no contact with young children. Dusty does not have assistance dog privileges in public, however she provides much needed companionship in the home and on countryside walks with her new handler.

Terry (a Cocker Spaniel from show lines) went from a score of 4 in test one to a score of 2 in test two. Terry had a good history of living and interacting calmly with junior aged children throughout socialising and training. On T1 he was tested with an unfamiliar 7 year old child greeting him in town. However on the second test the QAE, handler and dog were approached unexpectedly by a very boisterous 5 year old boy on the town part of the test. The child run toward Terry shouting and then grabbed Terry around the body unexpectedly. Terry showed strong fearful avoidance but no aggression. Unfortunately he then showed anxiety and avoidance meeting a 9 year old and a 3 year old on that same day. Due to Terry's impeccable history and the unusual situation, Terry was allowed to rest for one week and in the second week was gently and gradually re introduced to older and then younger children at the centre. The re-introduction was done on Terry's terms and was mostly either done on a long lead or off lead using positive associations between the child and food or his favourite toy. Terry was then re-tested two weeks later with an unfamiliar toddler and a junior aged child in town. On this re-test Terry remained relaxed while being stroked one at a time by each child. Terry has since gone out working as a full assistance dog and has had no further incidences.

3.3.6.2 Vocal reaction

Two dogs failed at T1 and then again at T2. Both were miniature Poodles and were brothers. The same QAE tested both dogs at the Southern centre for T1, however both dogs were re-tested at T2 by different QAE from the first and from each other.

One had their T2 at the Northern centre and one had their T2 at the Southern centre. Due to the vocals occurring in town situations and during sound work, neither dog could be placed as a full assistance dog or as a SSD dog. Both were later placed as CCD's.

Two dogs failed at T1 but then scored a 3 for their T2. Both dogs were female Cocker Spaniels from show lines. Neither dog was related to each other. Each dog had the same QAE for their first test as they did for their second. However different QAE's tested each dog. One dog had both tests at the Northern centre and the other dog had both tests and the Southern Centre. Both dogs were un-neutered, and had not had their first season by the time of their T1. Both dogs received continued training in this area between T1 and T2. On T1 both dogs vocals were occurring in the home, in shops and when settling in cafe type situations. It was noted that in nearly all occasions for both dogs, the vocals occurred when the dogs were frustrated or bored. Both dogs received positive training between T1 and T2 by different trainers. Both dogs also had a season and were later neutered during T1 and T2. Both dogs improved by their T2 and have since become full Assistance Hearing Dogs.

Three dogs got a 3 in their T1 and then a 2 in their T2. Dog 1 and 2 are cocker spaniels from working lines, One is a male (Tom) and one is a female (Teena). They are from the same litter. The third dog is a female poodle (Puddle). At the time of T1 all dogs were un neutered. The two females had not had a season yet. Teena was tested by the same QAE for T1 and T2. Teena had her T2 at the Southern Centre. Tom had his sound work tested at the socialiser's home but was tested in the same town as Teena was. Tom was tested by a different QAE for both T1 and T2. The person who tested Tom on his T2 was also the same person that tested Tina on her T2. Both dogs were neutered by the time of their T2. Puddle was tested at the Northern Centre for both T1 and T2. She was unneutered for her first test but not for her second. She was tested by the same QAE for both tests. Tom and Teena were vocally reacting towards novel objects in town and park environments. Puddle became vocally reactive towards novel people and during sound work.

By T2 Toms scores had declined from a pass in T1 to a fail in T2 in the further areas of 'Adaptability', 'Environmental', 'recovery rate', 'frustration' and 'trainability'. Teena's scores only failed in the further area of 'Adaptability'. Due to Tom's change in behaviour since coming into the training centre, Tom's training was moved to a

new socialiser's house after 4 weeks. He could not go back to the original socialiser as she lived too far from any of the training centres. From then onward all of Tom's sound work training was done from the new socialiser's home. Teena remained at the Hearing Dog Southern centre during the week days. However she went home to a socialiser at evenings and weekends.

Due to Toms poor scores declining in most areas after his T2 he had a change of career and became an SSD dog. Teena unfortunately got kicked by a volunteers horse 5 days after her T2 and broke her leg. Tina did eventually make a full recovery with veterinary care. However in the 6 months it took to heal staff noticed that her reactive barking had decreased, her submissive urinating had stopped and her overall confidence had increased. After her leg had healed she was given another test and all of her scores apart from 'Adaptability' had returned to exactly what they had been on her T1. Tara is now out working as a full Assistance Hearing Dog.

Puddle also failed in the area of 'Adults' (she had previously passed adults on T1) due to her reactive barking at people. Puddle had a change of career and became a CCD dog.

3.3.6.3. Chase

One dog (Riley) failed on T1 but passed on T2. He scored a 2 on his T1 and scored a 3 on his T2. This was a male cocker cross miniature poodle. He was not neutered on his fist test but was by his second. He had a different QAE for T1 and T2 but both tests were conducted at the Southern Training centre, although in different parks and in different towns.

After reviewing the QAE's notes on T1, it seems the score of 2 was given because at the time of T1 the socialisers reported that the dog was being very persistent chasing the socialiser's local sheep. On a recent holiday to the coast, he had also chased and swam out to sea after seagull that was flying over head. The socialisers were also having trouble with Riley chasing their pet guinea pigs around the house. They believed that if left unattended he would kill one. Riley responded to recall commands, whistle commands and 'leave it' commands apart from when he was chasing something. No chase was observed on the actual T1, although apparently Riley did not have any opportunity to chase anything, due to low wildlife at the time of the test.

Between T1 and T2 Riley continued training with the socialiser and allocated member of staff on his chase behaviour, although after T1 the frequency of the visits increased from once a month (as well as weekly puppy classes) to once every two weeks.

On T2 Riley was tested at the standard Southern nature reserve that all Southern dogs are tested at on their T2. Here he showed interest in a squirrel and some low flying bird, however he recalled away from both. He was not tested near sheep, seagulls or guinea pigs. Riley is now out working as a Full Assistance Hearing dog. He is mostly walked in recreational park type environments and rarely comes into contact with farm animals, sea gulls or guinea pigs.

3.4 Discussion

Many published studies (Wilson and Sungren 1996, Christensen et al 2006, Svobodova et al 2007, Weiss 2014, Guyer et al 2011) have explored behaviour tests on dogs or behaviour tests compared with questionnaires (Svartberg 2002,2005, Bollen and Horowitz 2008, De Meester et al 2011, Barnard et al 2012) However very few have any correlation with success in later performance using both passed and failed dogs. This study has been the first to evaluate, if an assistance hearing dogs performance in a later behaviour test, can be predicted from an earlier behaviour test. No dogs were excluded early on, so this work is the first to attempt to truly measure Hearing Dogs performance during training.

The hypothesis of this study was: Performance in the 18 month Hearing dog assessment can be predicted from the dog's performance in their 8 month hearing dog assessment. The research objectives from this initial exploratory analysis were to ascertain if the hearing dog assessments are able to show predictive validity with respect to task performance. With this in mind this work explored various measures of reliability and validity of the Hearing dogs behaviour tests currently in use.

3.4.1 Test for Social Behaviour (towards adults)

A significant but weak positive correlation was found between Hearing Dog social behaviour (adults) between T1 and T2 (See Table 6, fig 3). There was also no

significant difference in the performance of dogs at these times (Wilcoxon matched pairs test)(see Table 7) . Together these results would appear to indicate that the 8 month tests are predictive of later performance, and that there is no change in the level of performance over this time. However the weak correlation is seemingly due to the 14 dogs who's scores declined between T1 and T2, with 3 dogs actually failing their T2. (see fig 3)

Furthermore the sensitivity and specificity of the initial test suggested that the initial social behavioural test is highly sensitive to later performance (see Table 8), i.e. those that pass at an early age, also pass the later test (although this may partly be an artefact of the high success rate in the final test). By contrast the specificity level for the initial social behaviour test was extremely low (see Table 8), meaning that the initial tests ability to correctly predict a dog that would not pass was poor. After calculating the more refined measure of likelihood ratios (See Table 8), it is clear the Hearing Dog behaviour test for social behaviour (adult) has little predictive value for the 18 month test for social behaviour (adult)

When examining the dogs who failed their T2 who had previously passed their T1, it can be observed that two dog's obtained a 3 on T1, but this later dropped to a score of 2 by T2; one dog obtained a 4 on their T1 but this later dropped to a score of 2 on their T2.(see Fig 3) All dogs were male and both tests were conducted at the same centre for both T1 and T2. All dogs were un-neutered at T1 but neutered by T2. Each dog were assigned a different QAE from each other for their T1, however all dogs had the same QAE as each other for their T2. Although thorough training is given to the QAE, it could be speculated that the QAE who graded their T2 was perhaps a harsher marker. However the same QAE also lowered the dog's scores for T2 even though they previously assessed the same dog on the T1.

Interestingly all three dogs also failed their T2 for 'adaptability' for T1 and T2 (see fig 3) suggesting that it may have been the dogs overall poor ability to cope with the change of coming into advanced training from socialising, that affected their confidence with people. Two of the dogs' scores also declined in the area of children and vocals, suggesting that the dogs' poor ability to cope with change, may have had a knock on affect and manifested in other inappropriate social behaviours.

As previously discussed in the literature review, the confounding variables of the importance of the dog being leashed or not leashed, and gender of the stooge is an important factor that needs to be considered. The Hearing Dog 'Adult' T1 and T2 tests the dogs in multiple environments, on and off lead, with men and women. The history of the dog is also taken into account in all these areas. There is a note section when scoring for the QAE to mark down the dogs reactions with the different variables. However the score itself is a general score for all. For example to obtain a grade of 5 the dog must be 'confident to meet people'. It is unclear how the QAE should score if the dog is always fine with men, but occasionally scared of women when on lead etc. Greater clarification is needed to prevent ambiguity, if the T1 is to be a better predictor of T2.

For greater precision in general a larger sample size would have been beneficial, increasing the number of positive and negative results in each test, so the confidence interval around the sensitivity and specificity could have been more precise (Strassle et al, 2012). With so few Hearing dogs in this sample failing, this could potentially have yielded greater uncertainty in the sensitivity and specificity results.

For example in the test for social behaviour with adults, only one dog failed at T1 (with a score of 2) however he then went on to pass (with a score of 4) (see fig3). There are a number of different variables that could have played a part in this dog's scores changing between T1 and T2, and thus making it difficult for T1 to be a reliable predictor of later performance. The dog initially failed for being 'head shy with strangers and exhibiting hard mouthing behaviour and pulling at clothing in the home and in town with strangers. The dog was an adolescent male Labrador. On T1 he who un-neutered, but by T2 he was neutered and 10 months older. The hormonal change and increase in maturity may alone have corrected this behaviour.

However as soon as the dog completed his T1, his trainer also increased his home visits from once monthly to two weekly, with continued positive training between T1 and T2, which may have presented another confounding variable.

As a further measure of the quality of the test, convergent validity with relevant domains of the CBARQ questionnaire (Serpell and Hsu 2001) was assessed. CBARQ-1 correlated with the T1 for the area of stranger directed fear, but not for owner or stranger directed aggression (see Table 9). However CBARQ -2 correlated

with T2 for stranger directed aggression. Both correlations were negative as expected, due to the different scoring systems between the tests: a high score in the CBARQ meant the dog scored poorly, whereas a low score in the hearing dog test also meant the dog scored poorly. To further analyse predictive validity when just using the CBARQ questionnaire, the CBARQ-1 scores were compared with CBARQ-2. This showed a significant correlation between the 8 month and 18 month scores for stranger directed aggression but not for owner directed aggression, or stranger directed fear.

However CBARQ results also need to be interpreted with caution, due to the poor response rate for questionnaires being returned. The sample size of dogs providing data was only 12 and this small sample size could have weakened the validity of the statistical measurement.

This difference between the two CBARQ questionnaires could also be due to differing levels of expertise between the people filling out the questionnaires. Although all socialisers are given support by Hearing Dogs for Deaf People and many are experienced, they often lack the degree of professional training that the trainers have undergone. Tami and Gallagher (2009) found that owners often showed difficulties when distinguishing between aggression and confidence, with submissiveness often being labelled friendliness and playfulness. Whereas Wan et al (2012) found that experience played a big part in the ability to perceive differencing emotions in dogs, and this was most pronounced in correctly identifying fearful dog behaviour. Westgarth (2015) reminds us “it is possible to be looking at the same data from different perspectives and seeing a different “truth.” Whereas Mills(2008) highlights the need for greater clarification between assessors and criteria, when distinguishing between a potentially protective versus a malfunctioned response. Greater clarification of scoring may therefore be beneficial for both the adult behavioural test and the CBARQ

The lack of a positive relationship between the CBARQ-1 and CBARQ-2 scores might also be due to fear continuing to develop into and past adolescence, but also subside, and become extinct as an animal gets older (Shechner et al 2014). Therefore the behaviour may have simply changed in the interim 10 months between questionnaires, like it did for the adolescent male Labrador that scored 2 for T1 and 4 for T2.

3.4.2 Test for Social behaviour (towards children)

A moderate positive correlation was found between the T1 and T2 behavioural tests for social behaviour with children (see Table 6, figure 4). There was no significant difference between the medians of the T1 and T2 tests (see Table 7). From this it is speculated that the 8 and 18 month rates of performance are similar.

However a parallel issue as the previous test, relating to sensitivity and specificity was found. Implying the 8 month test has little predictive value for the 18 month test for social behaviour (children). However as previously highlighted, low frequency of dogs failing may have affected the confidence interval.

There are many potentially confounding variables involved between T1 and T2, all of which may reduce the reliability of predictive value of the Hearing Dog 'child test'. From the previously discussed case study, it can be seen that two dogs failed the 'Children' T1, but only one of them went on to pass. Both had fairly similar test circumstances. For T1 both dogs had the same QAE, the same test centre, were of the same age, sex and had similar sexual status. Increased support was also given in their own homes, or surrounding area up to the age of 14 months, however this was done by different trainers. At 14 months confounding variable of the location and team of staff may have affected the results; one was moved up North to continue advanced training and one returned to the Southern Centre. Also on T2 both dogs were tested by a different QAE from each other and from the previous test, possibly resulting in a different interpretation of behaviours. By T2 the Labrador had improved but the cocker spaniel had not. The results of the statistical tests cannot provide us with causation; however the many confounding variables between the current T1 and T2 make the predictive reliability of T1 poor.

Furthermore when looking at the case study of the dogs that passed T1 for children but then failed T2, both dogs went from a score of 4 to 2 on their second test (see fig 4). As previously discussed, one dog had a negative incident shortly after T1 and one had one on the actual T2. Both dogs received positive behaviour modification, however the different ages that the traumatic event took place seemingly may have affected their recovery and thus the reliability of T1 to correctly predict T2. The

behaviour modification was also given by different trainers perhaps causing another confounding variable.

A further point to consider is that a dog's negative experience could propagate a negative bias towards one age range or gender of child over another. Although the QAE's are encouraged to write the age range of the children in the notes section along with their scoring, the actual scoring system does not take this in to consideration. If a dog is tested with toddler on T1 and fails, but on T2 is tested with a 9 year old and passes, or vica-versa, there becomes a very confused picture of whether the dog is suitable to be around children. This all of course negatively affects the reliability of the predictive value of T1 and T2. A possible solution to this would be to split the test into different age ranges.

Lastly only certain sections of the CBARQ could be used to potentially validate the Hearing Dog behavioural test. Out of 100 CBARQ questions, two (11 and 37- (as seen in appendices 2) question a dog's response to children. As the internal consistency of the CBARQ sections have previously been found to be satisfactory and valid (Svartberg 2005) any cherry picking of individual questions from the sections may have compromised the overall validity of the questionnaire (Biljon, 2014). For this reason it was decided not to use the CBARQ to test for convergent validity for the behavioural test for children.

3.4.3 Test for Social behaviour (towards dogs)

A weak but significant positive correlation between the 8 and 18 month tests for social behaviour towards dogs was found (see Table 6, fig 5). There was no significant difference in the medians between the T1 and T2 Hearing Dog behavioural tests for social behaviour (dogs) (see Table 7).

Although a significant correlation was observed, one reason why it may be weak could partly be due to one dog failing its T1 and 6 dog's failing their T2. As a result the sensitivity and specificity measures, and the likelihood ratios also demonstrate a similar finding to the previous sensitivity and specificity tests (see Table 8). The specificity level for the initial social behaviour test was extremely low (see Table 8). Seemingly this is due to this tests poor ability to correctly predict a dog that would not

pass, for example 6 dogs failed their T2 who had passed their T1 (5 of these dogs previously scored a 3 and 1 previously scored a 4) (see fig 5).

With further examination into these dogs, it appears that all 6 dogs developed fear related reactive behaviour towards other dogs, which had previously not be reported on their T1. 5 of these dogs only failed in the behaviour factor of 'dogs' and in no other areas. One dog failed in nearly all behavioural factors by T2. All 6 dog's reactivity was first noted by the volunteers and staff members in the interim period between T1 and T2. Interestingly 5 of these dogs had the same QAE as each other for their first test, however only two of them had the same QAE for their T2. It could be that fearful reactivity developed through maturity after T1, or it could be that tester error on their T1 accounted for the difference in scores.

Due to there not being a perfect RHO score of 1, it seemingly suggests that the QAE are not biased by the previous scores, otherwise there would have been the exact same number of dogs passing and failing on T2 as T1. However there could be tester bias in other areas. For example the QAE's review the dog's history when scoring both T1 and T2; this lack of double-blind conditions may reduce objectivity when scoring any subsequent tests. If there had not been any negative comments on the dogs interactions with other dogs prior to T1, this may have pre disposed the QAE on the T1 to mark the dogs more favourably around other dogs. As dog to dog body language can be dependent on the conspecific they meet (Bradshaw et al, 2015) the QAE on T1 may have been biased to grade more favourably if no negative dog interactions happened that day. However by T2, all dogs had reports in their file of them reacting negatively around other dogs. Thus perhaps biasing the QAE on T2, before they had even seen the dog with other dogs.

Although all four QAE undergo thorough training, their intra-rater reliability ranged from only 62% to 85%. While their degree of inter- rater reliability agreement was 81% . When the QAE's undergo training they use stooge dogs that the QAE have very little prior knowledge of. The QAE do not read the stooge dog's history. However in every day practice the QAEs nearly always have prior knowledge of the dogs that they are assessing. This means that the organisations training is not equipping the QAE on how to grade correctly, using the current Hearing Dog grading system, which always takes history into account. The training is therefore not a reflection of every day working practise, and may be one of the reasons why the

reliability of the predictive value of temperament tests carried out on Assistance Hearing Dogs is poor.

In the literature review the importance of noting the conspecifics breed, sexual status, age, height and sex was highlighted. It was noted that very few peer reviewed research currently does this. Likewise the Hearing Dog scoring system does not currently reflect any information regarding the other dog which could also affect the ability of T1 to reliably predict the results of T2. Taking the dogs entire history into account will of course help test the dog with a wider range of dogs, however unless this information is properly noted on interactions, the information is too vague for the QAE to make an informed decision on.

When comparing the T1 'dog' test with the CBARQ-2 sections of: 'dog-directed aggression', dog directed fear and 'dog-rivalry', zero significant correlations were found. The same was found when comparing the T2 with the same concomitant 18 month CBARQ sections (see Table 9).

The CBARQ and the behaviour test may not have correlated because they are measuring on different time scales (Proyer 2007). The CBARQ questionnaire is measuring generalised responses and is purely retrospective in nature, whereas the Hearing Dog behaviour test focuses on the immediate behaviour as well as the dog's history. The questionnaire may therefore be reflective of the reporter's memory and feelings rather than the actual behaviour and the reality of the situation at the time. It is not surprising therefore that the two different methods do not converge.

3.4.4 Test for Environmental behaviour

This test purports to evaluate a dog's confidence levels in public environments. A weak correlation was found between T1 and T2 for the 'environmental' behavioural test.(See Table 6, fig 6,). This correlation is partly due to the 13 dogs that scored 3 on both T1 and T2 and the 20 dogs that scored 4 on both T1 and T2, however there were in fact 17 dogs whose scores declined between T1 and T2 as opposed to only 11 dogs whose scores increased (see fig 6).

As previously there was no significant difference between the performances of dogs for this test (see Table 7) most probably due to the large amount of scores that tied between T1 and T2. When the sensitivity and specificity was considered for this test, as before, sensitivity was high and specificity was low (see Table 8), meaning the test had little predictive power for identifying dogs that would later fail.

When comparing the CBARQ section on 'non-social' fear with the corresponding age behaviour tests, there was no significant correlation between either T1 or T2, indicating no convergent validity (see Table 9).

When looking at the sample group it can be seen that 3 dogs failed their T2 after initially passing their T1. 2 dogs passed their T2 after previously failing their T1 (see fig 6). All dogs that failed their environmental test did so through a lack of confidence in a town environment; however it does not stipulate what part of town in particular they lost confidence around. For example was it traffic, or different floor surfaces in shops. This vagueness could be one possible reason behind why there is no convergent validity, and only a weak correlation and low specificity for this factor.

In a town environment the dog is bombarded by olfactory, visual, touch and auditory stimuli. Consequently it can be difficult to link the behaviour response with the correct stimulus that the dog is reacting to. The CBARQ questionnaire attempts to do this by splitting up the different components, asking: how the dogs behave towards traffic, novel items, noisy events, such as thunderstorms, items moving in the wind. In contrast T1 and T2 groups all of these elements and asks the scorer to give an overall score based on the dogs confidence in multiple environments. By grouping the response to all environmental stimuli in one section, it potentially makes the assignment of the scores less precise as to which aspect of the 'environment' is actually being scored at any given time. This ambiguity could affect the overall predictive reliability of T1.

The literature review discussed the importance of allowing dogs sufficient time to settle into an environment as emotional reserves can be over stretched (Rayment et al 2015). It was also found that neophopia may become an issue in novel environments or with novel objects. On T2 the dog has been trained in the environment it is tested in but only over a 10 week period. Most dogs should be fairly familiar with the environment after visiting it once a week for 10 weeks, however as 17 of the dogs

scores declined between T1 and T2 it may suggest longer time is needed in this environment to allow more time for the sensitive dogs to settle in. Interestingly the three dogs that failed their T2, had all previously scored a 3 on their T1. All three dogs also failed their adaptability at T2, suggesting that the dogs that failed their environmental test were struggling to adapt to the change of environments.

3.4.5 Test for Recovery rate

There was not a significant correlation between the Hearing Dog T1 and T2 tests for recovery rate (see Table 6, fig 7), which purports to test for the dog's level of resilience. When looking at the data for 'recovery rate' (fig 7) it can be seen that the dogs scores did not change in the same way over time. The data is offset by outliers which have enough influence to lower the correlation coefficient to non significant; 16 dogs scored 4 on their T1, but only 3 on their T2; 5 dogs scored 4 on their T1 but only 2 on their T2; 1 dog scored a 3 on the T1 but only a 2 on their T2 (see fig 7).

The Wilcoxon showed a significant decrease in the differences; with the success rates declining by the 18 month test (see Table 7). As previously discussed in the results section the 22 declining data points, including the 16 dogs who scored 4 on their first test, but only a 3 on their second (see 7), may explain the declining median in the ranked pairs of scores.

Similar issues as discussed previously relate to the sensitivity and specificity results obtained for this test (see Table 8). A comparison between the behaviour tests and the CBARQ was not applied due to a lack of a concomitant section in the CBARQ.

Although the statistical tests are not able to show causation, it may be the change of environment and handler which caused the 22 declining data points. Stress and depression has been found to lower an animal's resilience and recovery rate (Pfau and Russo 2015). Approximately 10 weeks prior to T2 the dogs are moved to one of the training canters. The decline in the T2 scores may be reflective of a change in the dogs mood due to their recent environmental and social changes, and thus their ability to bounce back to a ' normal' rate as they were previously able to at 8 months.

However what is a ‘normal rate’? The Hearing Dog definitions deem a high recovery rate as dog returning to normal with ‘5 seconds’ or a ‘short amount of time’. However this may be problematic to score without a prior accurate knowledge of what ‘normal’ is for that individual dog, as well as ambiguity over what time counts as a ‘short amount of time’. For example animals can cope in two ways; passive or active (Budzyńska 2014) and may use a variety of methods in order to recover, for example chewing (Helmreich et al, 2012). It could be that a passive dog is deemed as having a high recovery rate, when in reality they may still be struggling to cope, just be less demonstrative about it.

A further factor to consider is that a dog’s recovery rate is dependent on the emotions they are experiencing. An anxiety driven response may have a quicker recovery time than a habitual pathological fear or phobia (Ohl et al 2008). Mills (2008) similarly cautions that when considering a dog’s maladaptive behaviour and any treatment, a distinction should be made between a stimulus dependent response (state) and an underlying predisposition (trait). Based on this, it may be suggestive that there is a difficulty in scoring the Hearing dogs ‘recovery rate’ meaning that performance in the Hearing dog T2 cannot be predicted from the dog’s performance in their T1

The immediacy of the source of the perceived threat may also affect a dog’s recovery rate (Ohio et al 2008) a variable which is not discussed or accounted for on the Hearing Dog behavioural test. It is impossible for the test to specify the specifics of every threat; therefore, the predictability of having one scoring system for all seems to suggest why there is low predictive validity for this test.

3.4.6 Test for Adaptability

The T1 and T2 behaviour test for ‘adaptability’ showed a positive moderate correlation (see Table 6, fig 8). As has been previously discussed , 5 dogs obtained a score of 2 on their T1, but then a score of 3 in the T2. Similarly 2 dogs scored a 3 on their T1 but then scored a 4 on their T2(see fig 8). These data points along with the data of 14 dogs scoring 3 on both their T1 and T2, and 15 scoring 4 on both their T1 and T2 (see fig 8) would have caused an overall positive correlation.

However the Wilcoxon test showed there was also a significance difference between paired test scores, showing dogs scores significantly declined (see Table 7). The Spearman correlation shows a monotonic relationship, meaning the variables move in the same relative direction, but not necessarily at the same constant rate. Therefore it is possible to have a positive correlation (trend) while simultaneously obtaining a significant negative difference between ranked paired data. As previously discussed this is due to four points on the axes and 22 dogs where their tests declined, importantly with some of them declining by two units (see fig 8). Thus increasing the difference between the medians in the paired data sets

The Sensitivity and specificity measures, as with all previous tests, showed high sensitivity but low specificity (see Table 8). However the specificity for 'adaptability' was slightly higher than other tests at 25%, possibly as a result of the 3 dogs that were predicted to fail and did indeed fail, although with a positive likelihood of 1.2 and negative likelihood ratio of 0.4 this seemingly still shows poor predictive validity for future failures.

The CBARQ section of 'attachment and attention seeking', which included questions relating to over attachment from the dog towards its handler, completed at 8 and 18 months (see Table 9) did not correlate with the equivalent aged behaviour tests, showing a lack of convergence. However when comparing the differences between CBARQ-1 and CBARQ-2 questionnaire for the factor of 'attachment and attention seeking', the results showed that there was a correlation over time. (see Table 9)

As previously discussed, between T1 and T2 the dogs' move from their full time socialiser's home into training and generally into another volunteers home. The 18 month test occurs approximately 10 weeks after this significant change to the dog's lifestyle, home, and familiar handler. Changes in routine and the absence of a familiar bond has been shown to have a profound effect on a dogs emotional homeostasis (Demeester et al 2008). As with 'recovery rate', seemingly the poorer 18 month Hearing dog, and CBARQ scores may be reflective of this heightened emotional state due to the significant change of lifestyle the dogs' has just undergone.

The differences found between the 8 and 18 month results, may also be due to operational and grading differences between the 8 and 18 month tests. The T1 involves the dog working alone with the QAE for a short period. The T1 score is based on how the dog performs with the QAE, as well as their history. In contrast the T2 test does not test the dog alone with the QAE at any point. The T2 test for 'adaptability' and the CBARQ score 'attachment and attention seeking' only scores retrospectively on 18 months' worth of history. This seems to suggest that T1 is a reflection of the dogs' performance on the day, whereas T2 is primarily functioning as an 18 month review not a behaviour test.

3.4.7 Test for Vocal reactivity

The scores in this thesis for vocal reactivity demonstrated a moderate positive correlation (see Table 6, fig 9), with no significance difference between the medians of the paired data. Seemingly suggest the levels of vocal reactivity stayed fairly consistent within the sample group.

As has been previously discussed in the case study and will be discussed later in this section, only 17 dogs out of 62 scores declined between T1 and T2, and only one of these dogs declined by 2 units (see fig 8). 6 dogs failed in total, with only 4 of these dogs having previously passed in their T1.

The sensitivity and specificity test suggests the test is sensitive to later performance but with a low specificity level (33%) (see Table 8) meaning that the 8 month test still has a low predictive ability. The specificity level was the highest of the specificity scores for the different behavioural factors. However due to only predicting that two dogs would fail their T2 when in fact 6 dogs failed, it was still suggestive of poor specificity. The likelihood ratios (see Table 8) did not alter the low specificity levels. The 8 month Hearing Dog behaviour test for vocal reactivity has little predictive value for the 18 month test.

A set of quantitative scores for the behavioural category of vocals was not available on the CBARQ Questionnaire; subsequently convergent validity could not be tested using the CBARQ.

Vocalisations in dogs can be for many purposes: alerting to intruders, herding livestock, detecting a scent trail (Mills and Zulch, 2010) when displaying separation anxiety (Yin and McCowan, 2004, Storengen et al 2014), attention seeking (Yin and McCowan, 2004), barking at unfamiliar people dogs or items (Péter et al, 2014) , in play with other dogs (Yin and McCowan, 2004), or as an aggressive warning (Horváth et al, 2007). However not all of these contexts are unwanted behaviour and not all are problematic to a handler. Yet the Hearing dog behaviour test generalises and groups all vocals as undesirable. As we can see from the case studies; the two dogs that failed at T1 but then scored a 3 for their T2 were vocalising out of boredom and frustration. Whereas the three dogs that scored a 3 in their T1 and than a 2 were either vocally reacting towards novel objects in town and park environments during training, or being vocally reactive towards novel people during town and sound work.

The Hearing Dog scoring definitions for 'vocal reactivity' requests the QAE to grade the frequency of vocal reactions from 'none', to ' frequent' , and the context of the situation from: 'Any' to 'most situations/environments' (See page 23). Yet it does not account for duration or intensity. The intensity and duration of a dogs bark can be a behavioural problem especially when in suburban areas (Yeon 2007). This may be especially problematic in Hearing dogs as a deaf recipient may be unaware of any noise pollution caused by their dog.

Barking can be caused by transient states such as arousal or motivation (Taylor et al 2014) and can be altered through learning and training (Mills and Zulch, 2010) such as was the case for the two case study dogs that failed on their T1 but passed on their T2 (see page 69), and especially the case of Teena (see page 69). Teena was given extra time to settle in due to a fractured leg, during this time, her confidence improved and her vocals decreased. This case study may seemingly show that for some dogs the standard 16 weeks of advanced training time may not provide every dog with enough time to settle in. Or perhaps instead it shows Hearing Dogs may benefit from increased maturity, before they are brought in to begin their advanced training.

3.4.8 Test for Motivation

The Hearing Dog 'motivation' definitions are based on the dog's response to social interaction and level of reward. Out of a sample of 62 dogs, only 1 failed the test for motivation on the T1 and zero failed at 18 months. The dog that did fail his T1, later scored a 3 for his T2 after continued exposure to toys and games in the interim time between T1 and T2 (see fig 11).

There was a positive but weak correlation between T1 and T2 (see Table 6, fig 11), and no significant differences found (see Table 6). The weak correlation may partly be due to the outlier of 8 dogs scoring 4 on T1, but only 3 on T2 (see fig 11). The CBARQ did not correspond for this test and was therefore not used.

The sensitivity and specificity measures with the added likelihood ratios suggested the 8 month test provided little predictive merit (see table 8), although this may partly be because of the low frequency of dogs failing this test.

In the T1 the QAE plays with a toy in the novel room to find out the dogs responsiveness and thus motivation toward an item. As has previously been discussed in the literature review, any novel handler performing a motivation test with a dog, may present a confounding variable (Rooney et al 2001, Bradshaw, 2010). How the dog plays with the toy may be more as a response to the handlers' play bows and level of interaction and excitement, than their actual interest in the toy (Rooney et al 2001, Bradshaw, 2010). Perhaps suggesting why only one dog out of 62 failed their T1.

Similarly Shimabukuro et al (2015) reminds us that a lack of familiarity with a handler may have an impact on the dog's level of motivation. Also Duranton and Gaunet (2015) propose that dogs' adjust their behaviour according to the bond they have with the handler. Based on this, it would be expected that more dogs should have failed their T1. However the T2 has a trainer working with the dogs as opposed to their more familiar socialiser, so perhaps it does explain why 14 of the dogs scores declined between T1 and T2 (See fig 11). Perhaps the 10 weeks of working with the handler is enough time for the majority of dogs to form a good bond with their handler (reflected in the overall positive correlation), however for more sensitive dogs more time may be needed.

3.4.9 Test for Trainability

The two Hearing Dog 'trainability' tests showed a significant positive but weak correlation (see Table 6, fig 12). The scores also demonstrated a significant decline on the Wilcoxon matched pairs test, with the scores decreasing by the T2 test. As previously discussed in the results only 61 dogs instead of 62 dogs were able to have their ranked scores analysed for 'Trainability'. This was due to one dog's scores being unmarked for his T2. It is unclear why this was done, however in reality it had no overall effect on the dog's pass or fail rate as a Hearing Dog. However this of course does need to be taken into consideration when reviewing the results, as the sample was smaller than the other behaviour tests, and may have skewed the data slightly.

As has been previously discussed in the results, the significant negative Wilcoxon was partly obtained due to 21 dogs ranks decreasing from T1 to T2, as opposed to only 4 dogs whose ranks increased from T1 to T2 (see fig 12) . 36 dogs scored both the same on their T1 as they did on their T2, helping to provide the overall positive but weak correlation (see fig 12), but overall negative difference. Interestingly 3 dogs dropped by 2 scores between T1 and T2, while 1 dog dropped 4 wholes scores between T1 and T2 (see fig 12). Both dogs that failed in T2 had previously obtained a score of 4 (see fig 12).

Zero dogs failed this test at 8 months old, but 2 failed at 18 months old (see fig 12). Due to the sensitivity and specificity being calculated on pass and fail rates, this in-turn demonstrates zero predictive validity and therefore yielded a 100% Sensitivity for this but only a specificity of 0% for this test (see Table 8).

The T1 compared with the CBARQ-2 scores showed no correlation, neither did CBARQ-1 when compared with CBARQ-2 scores. (see Table 9)

Interestingly when comparing the CBARQ-1 questionnaire which were filled out by the Hearing dog socialisers at 8 months old, compared with the CBARQ-2 questionnaire which were completed by the trainers at approximately 18 months old (see Table 9), they showed the same as the hearing dog behaviour tests; they showed the scores had also declined, and therefore convergent validity for this test.

The Hearing Dog trainability test is measured on the dog's ability to learn new behaviours. Because of this T2's test is different from T1's test. On T1 the dog is taught to target a cooker timer for the first time. On T2 instead the scores are based on the dog's recent ability to learn advanced sound work. There is a broad difference between the difficulty levels of these two tasks. Targeting a cooker time is a short sequence of events, whereas alerting to a sound is complex sequence and chain of tasks. Hearing Dogs made the second test harder to reflect the 10 months of interim training that has occurred since T1. However this does mean that T1 and T2 are no longer comparable due to their differing levels of complexity and completely different tasks.

However despite this fact 36 dogs training ability was graded the same between T1 and T2. To obtain a grade of 5 a dog has to "work consistently and will persevere despite significant increases in criteria expectation and or sudden changes in criteria within a training period" (Hearing Dogs For Deaf people, 2015). This definition does not detail the dogs ability to problem solve, but instead it describes the motivation of the dog to carry on despite criteria expectation. In fact when looking at the individual dogs it was observed that the two dogs that failed their T2 in the areas of trainability also failed in 'motivation'. Seemingly suggesting the 'trainability score' is completely dependent on the motivation score, and is a test of motivation, not the ability to learn a new task.

The two dogs that failed also failed for 'environmental' and 'recovery rate', perhaps suggesting that these dogs adaptability and confidence had declined since coming into training and affected other scores. It stands to reason that adaptability, trainability and motivation would be so linked; a dog who is struggling to adapt will possibly be less motivated and therefore will also struggle to be trained.

A further reason why the correlation between trainability scores was significantly consistent can be seen when we re-examine findings from the critical review from the research by Braem and Mills (2010). Where by responsiveness to a command was found to be affected by verbal information preceding the command (Braem and Mills 2010). The socialisers and trainers are taught to train the dogs in the same way, primarily with hand signals instead of verbal chat. This perhaps highlights why 36 dogs were able to remain constant in their scores and were able to receive the same

score for their T1 as their T2. The majority of the dogs were able to generalise the hand signals and lessons from their previous training to their new advanced training.

3.4.10 Frustration

The 'frustration' test consists of measuring the dogs level of calm when something of value is withheld, the data showed a positive but weak correlation between the 8 and 18 month tests (see Table 6, Fig 10) and no significant difference was found across the medians (see Table 7). The weak correlations may partly be due to 31 dogs obtaining the same score for their T1 as their T2. There were 16 cases where dogs' scores declined between T1 and T2, and 15 where dogs' scores increased (see fig 10). The sensitivity and specificity measures showed similar results to previous measures (see Table 8). Consequently there was no practical predictive value for this test.

Emotional conflict (frustrate non reward) occurs when a reward is suppressed (Kuhne et al 2013). A frustrated response is therefore dependent on the animal having motivation towards obtaining the reward. As previously discussed, motivation can be affected by a number of variables, and different coping mechanisms will influence a dog's reaction on the test. Therefore the Hearing Dog behavioural test for 'frustration' may be adversely affected by the same confounding variables as the Hearing Dog test for 'Motivation'. Interestingly the dogs that failed their frustration scores on their T2, did also score highly in the area of motivation on their T2. Further work looking into the correlation between the two scores would be beneficial for future work.

Behavioural responses to frustration such as increased activity and vocalisation have been observed in dogs (Lund and Jørgensen, 1999). Rearing , vocalisations and conditioned avoidance have been observed in rats (Bentosela et al 2008), redirecting and displaying displacement activities have been seen in hens (Bentosela et al 2008), and a general increase in activity and oral manipulation occurs in pigs (Lewis 1999). However all of the above behavioural factors can also be linked to other emotions such as excitability, impulsivity and arousal (Rayment et al 2015). Ambiguity may therefore be present in the correct identification of a frustrated dog, and explain why there were 16 cases where dogs scores declined between T1 and T2, and 15 where a

dogs scores increased. Especially as T1 and T2 may have different QAE grading them and therefore perhaps subject to tester error.

3.4.11 Test for Chase

There was a positive moderate correlation for the behavioural 'chase' test (see Table 6, fig 13), with 10 dogs increasing their scores between T1 and T2, and 34 dogs that scored the same on T1 as they did on T2. However interestingly 18 dogs scores declined between T1 and T2, but there was no significant difference over time (see Table 7). As the sample is a relatively small sample the correlation was possibly affected by the 24 dogs that scored 4 on both their T1 and T2, as well as the 3 dogs that scored a 5 on both their T1 and T2. No dogs moved more than one unit of score between T1 and T2, explaining why there was perhaps no significant difference between paired data.

There was a low predictive value for this test due to a low specificity. There were also no correlations between CBARQ-1 and CBARQ- 2 questionnaires with the equivalent aged 'chase' behaviour test (see Table 9). The CBARQs and the behavioural tests should both be reflective of the level of prey drive in the dog, however there was no convergence between the CBARQ and the HD behaviour tests.

One operational variable between T1 and T2 is that they are tested in different parks (South) or parts of the canal (North). The different areas have a suspected difference in wildlife populations. As both tests are graded with history in mind there is a confounding variable of 'opportunity '. For example if the dog is socialised in areas of low wildlife, and then also tested on their T1 in an area of low wildlife, than the dogs T1 will more than likely receive a positive score, as they have had little opportunity to demonstrate chase. The fact that only 1 dog failed on their T1 seemingly may reflects this. As all dogs are tested in a higher wildlife area on T2, this would seemingly increase the dogs opportunity to chase. This is perhaps the reason why 18 dogs scores declined between T1 and T2 (see fig 13).

From the previous case study it can be observed, for the dog that scored 2 on his T1, his score was based entirely on history not what was seen on the test. The negative history was based on persistent chasing of sheep in fields, seagulls at the beach, and

Guinea pigs around the house. Had he not had the opportunity to do this behaviour, or if his history was not taken into account when scoring, than his T1 Scores would have been very different. As it was detected early, the increased visits and training between T1 and T2 could present a confounding variable, and thus affect the reliability of the predictive value of temperament tests carried out at hearing Dogs For Deaf People. One way around this confounding variable would be to carry out T1 and T2 ideally in the same environment, however if this is not possible, both should be in areas of moderate wildlife.

3.4.12 Test for Hunt

There was a moderate positive correlation between the tests and no differences were observed between T1 and T2 (see Table 6, fig 14). Out of all the behavioural factors 'hunt' had the highest frequency of dogs scoring a 5 on both their T1 and T2 (8 dogs). 38 dogs scored exactly the same for their T1 as they did for their T2. 12 dogs increased their score, and 12 dogs decreased their score between T1 and T2 (see fig 14). No dogs increased or decreased their scores by any more than one unit (explaining the lack of significant Wilcoxon). This suggests that the moderate positive correlation is mostly due to the 38 dogs that scored exactly the same for both tests, including the 8 dogs that scored 5 on both tests.

The data produced a 100% sensitivity 0% specificity (see Table 8). This is due to the zero fail rates within both tests. Convergence could not be tested for hunt using the CBARQ.

The Hearing Dog's socialisers and trainers hone their dogs tracking skills to be co-dependent with them by playing reward based 'find it games' from an early age. This training continues between T1 and T2. The training has the benefit of moulding a dog to behave as part of a team with its handler, instead of tracking independently away from them (Wells and Hepper, 2003). It may be this lack of encouragement to track independently from the handler that has the influence over the low fail rates and high pass rates. 37% of the sample have working lines within them, with 9 being pure working cocker spaniels, however no dogs failed on their T1 or T2 despite this. In fact 4 of the working cocker spaniels scored both a 4 for their T1 and their T2. A

dogs' ability to determine directionality of a scent is aided by training when it is young (Wells and Hepper, 2003) therefore the amount of training that the dog receives prior to tests and between tests may explain why 38 dogs were able to remain constant with their scores across tests.

As has previously been discussed in 'chase' a lack of opportunity may have biased the results on T1. However all dogs are tested in moderate wildlife areas in T2, and as no significant differences were found between T1 and T2, this did not seen to be as much of a factor with 'hunt', however it could have accounted for the 12 dogs scores that declined between T1 and T2.

When looking at the definitions for 'hunt' it seems that the dog is graded negatively for demonstrating a hunt drive, but positively if it does not. As co-operative tracking is encouraged, but independent tracking is seen as problematic, it seems the dogs level of independence while on a walk, as opposed to their level of hunt drive is actually what the test aims to find out. A change of terminology may be beneficial for this test.

3.4.13 Test for Distractibility

The Spearman's correlation showed there was no significant correlation between the Hearing Dog tests for distractibility (See Table 6, Fig 15). There was not a significant difference between the medians (see Table 7). 8 dogs failed the test for distractibility at 8 months (with a score of 2), 7 of these dogs later scored a 3 on their T2 and one went on to score a 4 on their T2. Overall 16 dogs scores declined between T1 and T2. 18 dogs scores increased between T1 and T2, and 28 scored the same for T1 and T2 (see fig 15). The sensitivity and specificity test showed there was 0 % specificity in this test (see Table 8). The CBARQ could not be comparable for this test. Due to the above tests, the predictive validity of this test is low.

The lack of any correlation in this test may be in part due to the indistinctness of the test definitions. The test is graded from 1-5 on a dogs' level of interest to 'environmental stimuli'. Yet it does not stipulate which environmental stimuli in

particular are included, nor does it define how a dog's level of interest is measured. For example is it measured on the dogs gaze, or if it sniffs or tries to scavenge. Distractibility is tested in the town, the cafe and the park; with such a wide variety of test environments and ambiguity within the definitions, it is not surprising that it may be difficult for a level of agreement to be reached between tests.

A dogs attention span can be dependent on: the bond between dog and handler (Mongillo et al,2010) the dogs age (Whelan, 2006), its gender (Goddard and Beilharz, 1983), it's emotional state (Passalacqua et al, 2013) and the environment it is in (Åkerberg et al 2012). These are all confounding variables that may affect the predictive validity of the 8 month test.

Chapter 4

Overall Discussion

Contents **Page**

4.1 Overall Discussion and recommendations

By researching existing studies on behaviour tests and examining predictive validity of assessments with respect to task performance on published papers, this thesis was firstly able to identify that out of 17 papers looking at performance, only two papers Valsecchi et al (2001) and Vas et al (2008) were able to show predictive validity of their behaviour tests by performing re- tests on both passed and failed dogs by direct observation in the dog's natural environment. The author is aware of no other studies involving working dogs that are able to show the same similar level of predictive validity with respect to task performance, leaving a gap in research in this area.

The results from the review combined with the Hearing Dog behaviour tests highlighted the strengths and limitations of different methods of behaviour testing and the difficulties in trying to obtain face validity while avoiding confounding variables. The review highlighted how behaviour tests that are conducted in unnatural settings, are often subject to order effect, unfamiliar handlers, unfamiliar settings and neophobia; often then generalising their findings to real life without first re testing the dogs in a real life setting, resulting in low validity.

Hearing Dogs UK conduct behavioural tests and re-tests on both passed and failed working dogs, in a natural environment by direct behavioural observation. However until now the reliability of the predictive value of the 8 month behaviour test, has never been tested. In order for the test to be used as a reliable measure for the organisation, it was highly important that the behaviour test's predictive validity was assessed. This theses set out to do that.

However from the results gained from the Hearing Dog behaviour tests, it can be debated whether the 8 month behavioural test is beneficial to Hearing dogs for deaf people, as performance in the 18 month Hearing dog assessment cannot be predicted from the dog's performance in their 8 month hearing dog assessment. The 8 month assessment is unable to show validity with respect to task performance.

The areas of 'distraction' and 'recovery rate' were the only two factors that did not show any correlations between 8 and 18 month tests, with 'Recovery rate' also showing a significant decline between tests. The test for 'adaptability' and trainability also showed a significant decline between the 8 and 18 month tests. All

behaviour tests as previously mentioned were not able to demonstrate predictive validity through sensitivity and specificity measures.

The CBARQ scores were convergent with Hearing Dog scores in the areas of; T1 'adaptability' and CBARQ-1 'attachment and attention seeking'; T2 'dog' test and CBARQ-2 'dog directed aggression; 'adult' T2 and CBARQ-1 'stranger directed aggression'; T2 for 'environmental behaviour' and CBARQ-1 for 'non-social fear'; T2 for 'adaptability' with CBARQ-1 for 'attachment and attention seeking' and the T2 test for 'trainability' and CBARQ-1 for 'trainability'. As previously discussed the confounding variables may have been different levels of expertise when scoring a dogs behaviour, measuring on different time scales when compared with a behaviour test and experimenter expectancy. In addition a larger quota of returned CBARQ questionnaires would have been an advantage. The low return rate seemed unusual and may be as a product of the questionnaire not being computer based and easier to send back.

In hindsight one possible solution for a greater response rate may have been to make the Questionnaire computer based, so the participants could simply email back the response instead of posting it, or scanning and then emailing. Another solution may have been to use a questionnaire that was shorter as perhaps 100 questions seemed too long for some people to fill out. In an ideal world requesting the return in person would have been beneficial, however due to differing geographical areas this was not always possible.

It can be disputed whether all the Hearing Dog behaviour tests are needed in order to provide correct detailed information that is relevant for a working Assistance Hearing Dog. As the areas of 'Recovery rate' 'distractibility' and 'trainability' all showed a significant decline as the dog matured these tests at 8 months in particular show little predictive merit. As was also discussed the definitions for trainability seem to be closely linked to those of motivation, so perhaps both tests are not needed.

As discussed other sub tests may also benefit from a change in the terminology and a closer inspection as to what the aim of the test is and how it is scored. In particular 'Environmental behaviour' may benefit from being split into two sections; one score for traffic and one score for behaviour in shops. Likewise 'Vocal reactivity' may benefit from a change of terminology to include duration as well as frequency.

The critical review highlighted the need for behaviour tests to demonstrate content/face validity in more realistic settings. It seems beneficial having the Hearing Dog test in a setting similar to the type of environment in which the dog will eventually work. Naturally occurring events and stimuli will happen on the test that are more likely to happen when the dog is out working. However one seemingly unavoidable limitation of a naturalistic test, is the knock on effect the many variables may have on the behaviour that is being observed. For example on the test of 'social behaviour adults' the dogs ability to be relaxed when interacting with a stranger, when it may already be nervous from the passing traffic, or may already be struggling to adapt, may alter its state compared to if s/he were being tested in a familiar home. Therefore the validity of a given score is questionable when it is affected by so many other variables.

A potential suggestion as to how to overcome this would be to split each behaviour test (with the exception of hunt and chase) into three separate locations; 1) When in a familiar home environment 2) When the dog is in a familiar town environment and 3) Novel natural setting. This may provide information on a dog's performance before any anxiety sets in, and which behaviours will be exhibited if a stressful state in a new environment is reached (Schoening and Bradshaw, 2006). Staff members already fill in qualitative, narrative data , however by up skilling staff further as to which salient points to observe out in the field, and have them working in collaborating with the QAE scoring at one of the centres, it may save money by reducing the QAE time expenditure and travelling costs.

This method may also benefit the wider community as it may provide an opportunity to develop additional research as to how much an unfamiliar environment affects specific states or traits in dogs. This information may be beneficial to other assistance dog organisations as well as rescue centres.

One of the major advantages of the Hearing Dog behaviour test, is it is graded with knowledge of the dog's history in mind. This requires a collation of information from other staff members about the dog's behaviour beforehand. This approach differs from other behaviour tests (Wilson and Sungren 1996, Christensen et al 2006, Svobodova et al 2007, Weiss 2014, Guyer et al 2011), and offers an excellent opportunity to gather knowledge of dog over a long period of time in one central location. In turn providing a more detailed picture of dogs' behaviour over time and

contexts. In an ordinary controlled one- off behaviour test this information may be difficult to ascertain due to time constraints. (Rayment et al 2015). However as discussed history may also in turn bias the scores, and in some cases cause confusion as to how much of the score should be reflective of the dogs history, and how much should be reflective of the observations on the day. The tests would benefit from a clear distinction and further clarification. Perhaps then T1 would show greater predictive value.

As a measure of predictive validity the test is poor, feasibly due to the confounding variables that have already been discussed. However if viewed as an early 8 month warning system to highlight problematic behaviours, then it may well be highly beneficial. Due to the scope of this thesis it was impossible to look at more dogs on a case by case basis, which may have exposed this to be the situation. However the dogs that were discussed did seem to improve after increased training was given between T1 and T2. As discussed briefly before, perhaps the tests demonstrated poor predictive validity because it is primarily used this way, and therefore once an unwanted behaviour is highlighted, the dog is given more support and additional training. However for this to be concluded it needs to be established that the 8 months tests are at least reliable, and that was not possible to assess in this study, but should form focus of future work.

This study may have benefitted from a larger sample size; a solution would have been to test over a longer time frame. Unfortunately due to the nature of this thesis, on this occasion this was not possible. However it is something that may be useful to explore in the future. A further critique of this study is that the CBARQ questionnaire used may not have been suitable for this study. Many of the behavioural factors did not match up with the HD behavioural factors, and those that did rarely converged.

A further area that arose within exploration of the case studies, was that in some cases a dog was given a further re-test after they failed their T2, if they had previously shown good history and passed their T1. It may have been beneficial for this evaluation to have explored these post T2 re-tests, and evaluated if the subsequent scores improved or declined. This may have given more clarity over the stability and reliability of the T1 test compared with the T2 test.

It may be beneficial to further explore the return rate of any of the sample group who were ultimately placed with a recipient, and how these related to the T1 and T2 test scores. This would allow further examination of the predictive validity of the tests with respect to working performance, and ultimately assist Hearing Dogs produce dogs in a more efficient cost effective manner.

4.2 CONCLUSION

In conclusion the predictive validity of the 8 month Hearing dog behavioural tests is questionable with regard to its ability to correctly predict a dog's potential performance. Never-the-less this research speculates that this is due to many variables, one being the influence of continued training and socialising between the 8 month and 18 month tests and the ability of the test to act as an early warning system against unwanted behaviours. This research has highlighted the benefits and limitations of a behaviour test in a natural setting and the benefit of continuing to monitor and work with dogs that may have otherwise been rejected at an earlier stage.

References

- Åkerberg, H., Wilsson, E., Sallander, M., Hedhammar, A., Lagerstedt, A., Larhammar, D., Meyerson B.. (2012). Test for personality characteristics in dogs used in research. *Journal of Veterinary Behavior: Clinical Applications and Research* . 7 (6), 327-338.
- Arata, S., Takeuchi, Y., and Mori, Y. (2011) Toward an early prediction of guide dog qualification. *Journal of Veterinary Behavior: Clinical Applications and Research*, 6(1), pp. 63
- Asher, L., Blythe, S., Roberts, R., Toothill, L., Craigon, P.J., Evans, K.M., Green, M.J., and England, G.C.W. (2013) A standardized behavior test for potential guide dog puppies: Methods and association with subsequent success in guide dog training. *Journal of Veterinary Behavior: Clinical Applications and Research,* (0).
- Audrestch, H.M., Whelan, C.T., Grice, D., Asher, L., England, G.C.W., Freeman, S.L. (2015). Recognizing the value of assistance dogs in society. *Disability and Health Journal* . 8 (4), 469-474.
- Barnard, S., Siracusa, C., Reisner, I., Valsecchi, P., and Serpell, J.A. (2012) Validity of model devices used to assess canine temperament in behavioral tests. *Applied Animal Behaviour Science*, 138(1–2), pp. 79-87.
- Batt, L., Batt, M., Baguley, J., and McGreevy, P. (2008) The effects of structured sessions for juvenile training and socialization on guide dog success and puppy-raiser participation. *Journal of Veterinary Behavior: Clinical Applications and Research*, 3(5), pp. 199-206.
- Batt, L. S., Batt, M. S., Baguley, J. A., & McGreevy, P. D. (2008). Factors associated with success in guide dog training. *Journal of Veterinary Behavior: Clinical Applications and Research*, *3*(4), 143-151.
- Berns, G.S.,Brooks, A.M and Spivak, M. (2015). Scent of the familiar: An fMRI study of canine brain responses to familiar and unfamiliar human and dog odors. *Behavioural processes,*110(0),pp. 37-46
- Bentosela,M., Barrera, G., Jakovcevic, A., Elgier,A.M., Mustaca, A.E.. (2008). Effect of reinforcement, reinforcer omission and extinction on a

communicative response in domestic dogs (Canis familiaris).*Behavioural Processes* . 78 (1), 464–469.

- Bian, Y., Pan, Z., Hou, Z., Huang, C., Li, W., Zhao, B., (2012). Learning, memory, and glial cell changes following recovery from chronic unpredictable stress. Brain Res. Bull. 88, 471–476.
- Biljon, J.V. (2014). Questioning the questionnaire: Expediency of reviewing and publication verses adequate description and methodological justification. In: Devos, J., De Haes, s. *8th European Conference on IS Management and Evaluation:* . Reading: Academic conferences and publishing international limited. 262-271.
- Blackwell, E.J., Bradshaw, J.W.S.,Casey. R.A., . (2013). Fear responses to noises in domestic dogs: Prevalence, risk factors and co-occurrence with other fear related behaviour. *Applied Animal Behaviour Science*. 145 (1-2), 15-25.
- Bollen, K.S. and Horowitz, J. (2008) Behavioral evaluation and demographic information in the assessment of aggressiveness in shelter dogs. *Applied Animal Behaviour Science*, 112(1–2), pp. 120-135
- Bradshaw, J. W., & Nott, H. M. (1995). Social and communication behaviour of companion dogs. *The domestic dog: Its evolution, behaviour and interactions with people*, *1*, 115-130.
- Bradshaw, J.W.S.. (2006). The Evolutionary Basis for the Feeding Behavior of Domestic Dogs (Canis familiaris) and Cats (Felis catus).*The journal of nutrition*. 136 (1), 1927S-1931S.
- Bradshaw, J.W.S., Pullen, A.J., Rooney, N.J. . (2015). Why do adult dogs 'play'?. *Behavioural Processes*. 110 (1), 82-87.
- Braem, M.D., Mills, D.S.. (2010). Factors affecting response of dogs to obedience instruction: A field and experimental study. *Applied Animal Behaviour Science*. 125 (1), 47–55.
- Braem, M.D., Mills, D.S.. (2010). Factors affecting response of dogs to obedience instruction: A field and experimental study. *Applied Animal Behaviour Science*. 125 (1), 47–55.
- Bridgman G, Macpherson B, Rako M, Campbell J, Manning V, Norman-Kelly T. (2000) A national epidemiological survey of mental illness in the New Zealand Deaf community. In: Hjortsö T, von der Lieth L, Carlsen C, eds.

Mental health services for deaf people: a worldwide perspective. Devon, UK: European Society for Mental Health and Deafness, 216–34

- Budzynska, M.. (2014). Stress Reactivity and Coping in Horse Adaptation to Environment. *Journal of Equine Veterinary Science*. 34 (1), 935–941.
- Christensen, E., Scarlett, J., Campagna, M., and Houpt, K.A. (2007) Aggressive behavior in adopted dogs that passed a temperament test. *Applied Animal Behaviour Science*, 106(1–3), pp. 85-95.
- Conley, M.J, Fisher, A.D, Hemsworth, P.H. (2014). Effects of human contact and toys on the fear responses to humans of shelter-housed dogs. *Applied Animal Behaviour Science*. 156 (1), 62-69.
- Costello, H., Holt, G., Cain, N., Bradley, E., Torr, J., Davis, R., Edwards, N., Lennox, N. G. and Weber, G. (2007). Professional training for those working with people with intellectual disabilities and mental health problems. In Bouras, N. and Holt, G. (Ed.), Psychiatric and Behavioural Disorders in Intellectual and Developmental Disabilities 2nd ed. (pp. 400-411) Cambridge: Cambridge University Press.
- Dammeyer J. (2010) Psychosocial development in a Danish population of children with cochlear implants and deaf and hard-of-hearing children. *Deaf Stud Deaf Education*. 15 (1), 50–58.
- De Meester, R. H., De Bacquer, D., Peremans, K., Vermeire, S., Planta, D. J., Coopman, F., & Audenaert, K. (2008). A preliminary study on the use of the Socially Acceptable Behavior test as a test for shyness/confidence in the temperament of dogs. *Journal of Veterinary Behavior: Clinical Applications and Research*, *3*(4), 161-170.
- Deldalle, S., & Gaunet, F. (2014). Effects of 2 training methods on stress-related behaviors of the dog (Canis familiaris) and on the dog–owner relationship. *Journal of Veterinary Behavior: Clinical Applications and Research*, *9*(2), 58-65.
- Deiderich,C. and Giffroy, J. (2006). Behavioural testing in dogs:A review of methodology in search of standardisation . *Applied Animal Behaviour Science*. 57 (1), 51-72.
- Diedrich, C. and Giffoy,J. (2009). Temperament testing, from puppies to adulthood. *Journal of Veterinary Behavior: Clinical Applications and Research*, 4(6), pp. 237-238.

- Duranton, C., Gaunet. F. . (2015). Canis sensitivus: affiliation and dogs' sensitivity to others' behavior as the basis for synchronization with humans?. *Journal of Veterinary Behavior*. – accepted but not printed manuscript
- Fatjo,J.,Feddersen-Petersen, D.,ruiz De La Torree Jose Luis.,Amat,M.,Mets,M.,Braus,B.andManteca,X. 2007. Ambivalent signals during agonistic interactions in a captive wolf pack. *Applied Animal Behaviour Science,* 105(4), pp. 274-283
- Feddersen-Petersen, D. U. (2007). Social behaviour of dogs and related canids.*The Behavioural Biology of Dogs. CAB International, Wallingford, UK*, 105-119.
- Fellinger J, Holzinger D, Dobner U. (2005) Mental distress and quality of life in a deaf population. *Soc Psychiatry Psychiatr Epidemiol* 40 (1) , 737–42.
- Fellinger,J, Holzinger, D, Pollard, R.. (2012). Mental health of deaf people. *Lancet 2012*. 379 (1), 1037–44.
- Finkelstein, J.C. (1976). Experimenter expectancy effects. *Nonverbal Communication*. 26 (3), 31-38.
- Forkman, B., Boissy, A., Meunier-Salaun M.C., Canali, E., Jones, R.B. (2007) A critical review of fear tests used on cattle, pigs, sheep poultry and horses. *Physiology & behaviour*, 92, pp. 340-374
- Fox, M.W. (1969). Behavioral Effects of Rearing Dogs With Cats During the 'Critical Period of Socialization'. *Behaviour*. 35 (3), 273-280.
- Fukuzawa, M., Mills, D.S, Cooper,J.J. (2005). More than just a word: non-semantic command variables affect obedience in the domestic dog (Canis familiaris). *Applied Animal Behaviour Science,* . 91 (1-2), 129-141.
- Goddard,M.E.and Beilharz, R.G.,1985. Individual variation in agonistic behaviour in dogs. *Animal Behaviour,* 33(4), pp. 1338-1342
- Goddard. M.E., Beilharz, R.G.. (1983). Genetics of traits which determine the suitability of dogs as guide-dogs for the blind. *Applied Animal Ethology*. 9 (3-4), 299-315.
- Guest, C. M., Collis, G. M., Mcnicholas, J. 2006. Hearing dogs: A longitudinal study of social and psychological effects on deaf and hard-of-hearing recipients. *Journal of Deaf Studies and Deaf Education*, 11, 252-261.

- Hall, S. L. (1998). Object play by adult animals. *Animal play: Evolutionary, comparative, and ecological perspectives*, 45-60.
- Hart L.A, Zasloff, R.L, Benfatto A.M. (1996). The socializing role of hearing dogs . *Applied Animal Behaviour Science* . 47 (1), 7-15.
- Hearing Dogs for Deaf People 2015. About our working dogs. http://www.hearingdogs.org.uk/what-we-do/our-dogs [Accessed 09.09/15)]
- Helmreich, D.L., Tylee, D., Christianson J, P., Kubala, K.H., Govindarajan, S.T., O'Neill, W.E., Becoats. K., Watkins, L., Maier, S.F. . (2012). Active behavioral coping alters the behavioral but not the endocrine response to stress. *Psychoneuroendocrinology* . 37 (1), 1941—1948.
- Heine, V.M., Maslam, S., Zareno, J., Joels, M., Lucassen, P.J., (2004). Suppressed proliferation and apoptotic changes in the rat dentate gyrus after acute and chronic stress are reversible. Eur. J. Neurosci. 19, 131–144
- Horváth,Z., Igyártó, B,Z., Magyar, A., Miklósi, A. . (2007). Three different coping styles in police dogs exposed to a short-term challenge. *Hormones and Behavior*. 52 (5), 621-630.
- Isgor, C Kabbaj, M. Akil, H. Watson, S.J. (2004). Delayed effects of chronic variable stress during peripubertal-juvenile period on hippocampal morphology and on cognitive and stress axis functions in rats. *Hippocampus*, 14 (1),636–648
- Kaulfuß, P., & Mills, D. S. (2008). Neophilia in domestic dogs (Canis familiaris) and its implication for studies of dog cognition. *Animal cognition*, *11*(3), 553-556.
- Kerepesi, A., Doka, A. and Miklosi, Á., 2015. Dogs and their human companions: The effect of familiarity on dog–human interactions. *Behavioural processes*, **110**(0), pp. 27-36.
- Kolibiki, H.M. (2014). A Study of emotional relationships among deaf adolescents. *Social and Behavioral Sciences*. 114 (1), 399-402.
- Kuhne, F., Sauerbrey,A.F.C., Adler, S. . (2013). The discrimination-learning task determines the kind of frustration-related behaviours in laying hens (Gallus gallus domesticus). *Applied Animal Behaviour Science*. 148 (3-4), 192–200.

- Lane, D. R., Mcnicholas, J., Collis, G. M. 1998. Dogs for the disabled: Benefits to recipients and welfare of the dog. Applied Animal Behaviour Science, 59, 49-60.
- Lefebvre, D., Diederich, C., Delcourt, M., & Giffroy, J. M. (2007). The quality of the relation between handler and military dogs influences efficiency and welfare of dogs. *Applied Animal Behaviour Science*, *104*(1), 49-60.
- Lewis, N.J. (1999). Frustration of goal-directed behaviour in swine. *Applied Animal Behaviour Science*. 64 (1), 19–29.
- Lucidi, P., Bernabò, N., Panunzi, M., Villa, P.D., and Mattioli, M. (2005) Ethotest: A new model to identify (shelter) dogs' skills as service animals or adoptable pets. *Applied Animal Behaviour Science*, 95(1-2), pp. 103-122.
- Lunda,J.D., Jørgensenb, M.C. . (1999). Behaviour patterns and time course of activity in dogs with separation problems. *Applied Animal Behaviour Science*. 63 (3), 219–236.
- Mathews, I.Z., Wilton, A., Styles, A., McCormick, C.M., (2008). Heightened neuroendocrine function in males to a heterotypic stressor and increased depressive behaviour in females after adolescent social stress in rats. Behav. Brain Res. 190, 33–40
- Meehan,C.L, Mench, J.A. (2002). Environmental enrichment affects the fear and exploratory responses to novelty of young Amazon parrots.*Applied Animal Behaviour Science*. 79 (1), 75-88.
- Mills, D. (2008). Recognising the nature of problem behaviour. *The Veterinary Journal . 176 (1), 127-128.*
- Mills, D, Zulch, H . (2010). Veterinary medicine and animal behaviour: Barking up the right tree!. *The Veterinary Journal* . 184 (1), 119–120.
- Mizukoshi, M., Kondo, M., and Nakamura, T. (2008) Evaluation of the potential suitability of guide dog candidates by continuous observation during training. *Journal of Veterinary Behavior: Clinical Applications and Research*, 3(5), pp. 193-198.
- McGarrity, M. E., Sinn, D. L., & Gosling, S. D. (2015). Which personality dimensions do puppy tests measure? A systematic procedure for categorizing behavioral assays. *Behavioural processes*, *110*, 117-124.

- Mongillo, P., Bono, G., Regolin, L., Marinelli, L.. (2010). Selective attention to humans in companion dogs, Canis familiaris. *Animal Behaviour* . 80 (6), 1057-1063.
- Murphey, Robert M., Francisco A. Moura Duarte, and Maria Cecília Torres Penedo.(1981) "Responses of cattle to humans in open spaces: Breed comparisons and approach-avoidance relationships." *Behavior Genetics* 11.1: 37-48.
- Murphy, J. (1995) Assessment of the Temperament of Potential Guide Dogs. Anthrozoos: *A Multidisciplinary Journal of the Interactions of People & Animals*, 8(4), pp. 224-228.
- Murphy, J.A. (1998) Describing categories of temperament in potential guide dogs for the blind. *Applied Animal Behaviour Science,* 58(1–2), pp. 163-178.
- Ohl, F., Arndt, S.S., van der Staay, F.J., (2008). Pathological anxiety in animals. The Veterinary Journal 175 (1), 18–26
- Ortolani, A., Vernooij, H. and Coppinger, R., 2009. Ethiopian village dogs: Behavioural responses to a stranger's approach. *Applied Animal Behaviour Science,* 119(3–4), pp. 210-218.
- Packard,J. Wolf behaviour:reproduction,social and intelligent: D. Mech, L. Boitani (Eds.) (2003) Wolves: Behavior, Ecology and Conservation, University of Chicago Press, Chicago (2003), pp. 35–65
- Page, M. M. (1973). On detecting demand awareness by postexperimental questionnaire. *The Journal of Social Psychology, 91,* 305-323.
- Passalacqua, C., Marshall-Pescini, S., Merola, I., Palestrini, C., Previde, E.P.. (2013). Different problem-solving strategies in dogs diagnosed with anxiety-related disorders and control dogs in an unsolvable task paradigm. *Applied Animal Behaviour Science*. 147 (1-2), 139-148.
- Péter,P., Éva,S., Anna, K., András, P., Ádám M.. (2014). More than noise?—Field investigations of intraspecific acoustic communication in dogs (Canis familiaris). *Applied Animal Behaviour Science*. 159 (1), 62-68.
- Pfau, M.L., Russo, S.J.. (2015). Peripheral and central mechanisms of stress resilience. *Neurobiology of Stress* . 1 (1), 66-79.
- Pongracz, P., Miklosi, A., Doka, A., Csanyi, V., (2003). Successful application of video-projected human images for signalling to dogs. Ethology 109, 809–821

- Proyer, R.T . (2007). Convergence of conventional and behavior-based measures:Towards a multimethod approach in the assessment of vocational interests. *Psychology Science*. 49 (2), 168-183.
- Ratcliffe,V.F., McComb,K.,and Reby,D.,2014. Cross-modal discrimination of human gender by domestic dogs. *Animal Behaviour,* **91**(0), pp. 127-135
- Rayment, D.J., De Groef,B., Peters, R.A., Marston, L.C. . (2015). Applied personality assessment in domestic dogs: Limitations and caveats. *Applied Animal Behaviour Science* . 163 (1), 1-18.
- Radley, J.J., Rocher, A.B., Janssen, W.G., Hof, P.R., McEwen, B.S., Morrison, J.H., (2005). Reversibility of apical dendritic retraction in the rat medial prefrontal cortex following repeated stress. Exp. Neurol. 196, 199–203
- Rintala, D. H., Sachs-Ericsson, N., Hart, K.A. 2002. The effect of service dogs on persons with mobility impairments: A pre-post study design. *SCI Psychosocial Process*, 15, 70-82.
- Rooney, N.J., Cowan S. (2011). Training methods and owner–dog interactions: Links with dog behaviour and learning ability. *Applied Animal Behaviour Science*. 132 (3-4), 169-177.
- Rooney, N. J., & Bradshaw, J. W. (2003). Links between play and dominance and attachment dimensions of dog-human relationships. *Journal of applied animal welfare science*, *6*(2), 67-94.
- Rooney, N. J., Bradshaw, J. W., & Robinson, I. H. (2001). Do dogs respond to play signals given by humans?. *Animal Behaviour*, *61*(4), 715-722.
- Rooney, N.J., Gaines, S.A. and Bradshaw, J.W., 2007. Behavioural and glucocorticoid responses of dogs (Canis familiaris) to kennelling: investigating mitigation of stress by prior habituation. *Physiology & Behavior*,*92*(5), pp.847-854.
- Sachs-Ericsson, N., Hansen, N. K., Fitzgerald, S. 2002. Benefits of assistance dogs: A review. *Rehabilitation Psychology*, 47, 251-277.
- Sandi, C., Davies, H.A., Cordero, I., Rodrigiz, J.J., Popv, V.I., Stewart, M.G., (2003). Rapid reversal of stress induced loss of synapses in CA3 of rat hippocampus following water maze training. Eur. J. Neurosci. 17, 2447–2456.
- Schoening, B., Bradshaw, J.W.S., 2006. A behavioural test of aggression for adult dogsdvalidity and welfare considerations. In: Heath, S., De Keuster, T.

(Eds.) Proceedings of the 12th European Congress on Companion Animal Behavioural Medicine. Ghent, Belgium.

- Seksel, K., Mazurski, E. J., & Taylor, A. (1999). Puppy socialisation programs: short and long term behavioural effects. *Applied Animal Behaviour Science*, *62*(4), 335-349.
- Serpell, J.A & Hsu, Y.A. (2005). Effects of breed, sex, and neuter status on trainability in dogs. *Anthrozoös: A multidisciplinary journal of the interactions of people and animals*. 18 (3), 196-207.
- Serpell, J.A. and Hsu, Y. (2001) Development and validation of a novel method for evaluating behavior and temperament in guide dogs. *Applied Animal Behaviour Science*, 72(4), pp. 347-364.
- Sforzini, E., Michelazzi, M., Spada, E., Ricci, C., Carenzi, C., Milani, S., Luzi, F., Verga, M. (2009). Evaluation of young and adult dogs' reactivity. *Journal of Veterinary Behavior: Clinical Applications and Research*. 4 (1), 3-10.
- Shechner,T. Hong, M, Brittonc, J.C, Pine, D, S.Fox N, A. (2014). Fear conditioningand extinction across development: Evidence from human studies and animal models. *Biological Psychology*. 100 (1), 1-12.
- Shimabukuro, C., Putrino,N., Helbling,J., Tognetti, S.,Bentosela, M.. (2015). Individual differences in social and non-social behaviors in domestic dogs (Canis familiaris) during the acquisition, extinction and reacquisition of a problem solving task. *Behavioural Processes* . 113 (1), 179–186.
- Sinn, D.L., Gosling, S.D., and Hilliard, S. (2010) Personality and performance in military working dogs: Reliability and predictive validity of behavioral tests. *Applied Animal Behaviour Science,* 127(1–2), pp. 51-65.
- Slabbert, J.M. and Odendaal, J.S.J. (1999) Early prediction of adult police dog efficiency—a longitudinal study. *Applied Animal Behaviour Science*, 64(4), pp. 269-288.
- Strassle, P. Hess, A.S. Thom, K.A. Harris, A.D. (2012). Assessing Sensitivity and Specificity in New Diagnostic Tests: The Importance and Challenges of Study Populations. *Infect Control Hosp Epidemiol*. 33 (11), 1177–1178.
- Stoehr, A, M. (1999). Are significance thresholds appropriate for the study of animal behaviour?. *Animal Behaviour*. 57 (5), 22–25.

- Storengena, L.M., Bogea,S.C.K., Strøma, S.J., Løbergb, G., Lingaas, F.. (2014). A descriptive study of 215 dogs diagnosed with separation anxiety. *Applied Animal Behaviour Science*. 159 (1), 82–89.
- Sümegi, Z., Oláh, K., Topal, J. . (2014). Emotional contagion in dogs as measured by change in cognitive task performance. *Applied Animal Behaviour Science*, . 160 (1), 106-115.
- Sueda, K.L., Malamed, R. (2014). Canine Aggression Toward People : A Guide for Practitioners. *Veterinary Clinics of North America: Small Animal Practice* . 44 (3), 599–628.
- Svartberg, K. (2002) Shyness–boldness predicts performance in working dogs. *Applied Animal Behaviour Science*, 79(2), pp. 157-174.
- Svartberg, K., Tapper, I., Temrin, H., Radesäter, T., and Thorman, S. (2005) Consistency of personality traits in dogs. *Animal Behaviour*, 69(2), pp. 283-291.
- Tami,G. Gallagher, A. (2009). Description of the behaviour of domestic dog (Canis familiaris) by experienced and inexperienced people.*Applied Animal Behaviour Science* . 120 (1), 159–169.
- Taylor, A.M., Ratcliffe, V.F., McComb, K., Reby, D.. (2014). Chapter 5 Auditory Communication in Domestic Dogs: Vocal Signalling in the Extended Social Environment of a Companion Animal. In: *The Social Dog. Behaviour and Cognition.*. School of Psychology, University of Sussex, Falmer, UK: Academic Press. 131-163.
- Tempany, C., Mills, D., (2008). Are dogs dummies when it comes to their perception of people?. University of Lincoln
- Tomkins, L.M., Thomson, P.C., and McGreevy, P.D. (2011) Lateralization and temperament tests can predict guide dog success. *Journal of Veterinary Behavior: Clinical Applications and Research,* 6(1), pp. 89.
- Valsecchi, P., Previde, E.P., Accorsi, P.A., and Fallani, G. (2010) Development of the attachment bond in guide dogs. *Applied Animal Behaviour Science*, 123(1–2), pp. 43-50.
- Valsecchi, P., barnard, S., Stefanini, C., and Normando, S. (2011) Temperament test for re-homed dogs validated through direct behavioural observations in shelter and home environment. *Journal of veterinary behaviour: Clinical Applications and Research,* 6(3), pp.161-177

- Vas, J., Topál, J., Gyori, B., and Miklósi, A. (2008) Consistency of dogs' reactions to threatening cues of an unfamiliar person. *Applied Animal Behaviour Science*, 112(3-4), pp. 331-344.
- Virányia, Z., Topálb, J,. Gácsib, M., Miklósia,A., Csányia, V. (2004). Dogs respond appropriately to cues of humans' attentional focus.*Behavioural Processes* . 66 (2), 161–172.
- Wan, M, Bolger, N, Champagne FA (2012) Human Perception of Fear in Dogs Varies According to Experience with Dogs. *PLoS ONE* 7(12)
- Wells, D.L.. and Hepper, P.G (1999) Male and female dogs respond differently to men and women. *Applied Animal Behaviour Science,* 61(4), pp. 341-349.
- Wells, D.L., Hepper, P.G.. (2003). Directional tracking in the domestic dog, Canis familiaris. *Applied Animal Behaviour Science* . 84 (1), 297–305.
- Westgarth, C.. (2015). Why nobody will ever agree about dominance in dogs. *Journal of Veterinary Behavior (2015) 1e3*. xxx (1), 1-3.
- Whelan, F. (2006). Socialite or social disaster?. *Veterinary Nursing Journal* . 21 (1), 16-17.
- Wilsson, E., and Sinn, D. (2012). Are there differences between behavioral measurement methods? A comparison of the predictive validity of two ratings methods in a working dog program. *Applied animal behaviour science*. 7 (141), p158-172.
- Wilsson, E. and Sundgren, P. (1997) The use of a behaviour test for the selection of dogs for service and breeding, I: Method of testing and evaluating test results in the adult dog, demands on different kinds of service dogs, sex and breed differences. *Applied Animal Behaviour Science*, 53(4), pp. 279-295.
- Yeon, S.C. . (2014). The vocal communication of canines. *Applied Animal Behaviour Science* . 159 (1), 82–89.
- Yin,S., McCowan, B.. (2004). Barking in domestic dogs: context specificity and individual identification Original Research Article. *Animal Behaviour*. 68 (2), 343-355.

Appendices

Contents Page

Hearing Dogs Dog Standards Training Document

This document is a guide to the standards of dog behaviour within Hearing Dogs for Deaf People. It should be used **only** as a staff training document and applied in combination with the 'Policy for Social Behaviour'.

The areas included in this document are:

- Social Confidence
- Environmental Confidence

Each area has a list of behaviours which have been classified as either 'acceptable', 'undesirable' or 'depends'.

Behaviours within the areas of social and environmental confidence have been divided into behaviour level 1 and 2. These levels have been developed to work in accordance with the temperament assessment throughout supply, socialising, training and partnership. Dogs categorised as showing behaviour level 1 will generally be marked as a 2 on the temperament assessment and those showing behaviour level 2 will be marked as a 1.

Dogs should be assessed individually with careful consideration allowing for:

- Age
- Stage of Training
- Frequency and cause of Behaviour
- Risk Assessment

For example, a dog in Puppy Socialising may not be considered unsuitable due to 'sensitivities' however; a similar dog, post training/behavioural work, may be considered unsuitable in the later stages of training.

If a dog, placed with a client, develops a behaviour that is considered to be undesirable, it may be considered appropriate to continue with the partnership and provide adequate support.

Social Behaviour

Level 1 **Level 2**

Behaviour with men, women, children, dogs and other animals

Frequent jumping up on greeting

Acceptable☐ Undesirable☐ Depends☒

Mouthing on greeting

Acceptable☐ Undesirable☒ Depends☐

Vocal spooking

Acceptable☐ Undesirable☐ Depends☒

Avoidance

Acceptable☐ Undesirable☐ Depends☒

Barking

Acceptable☐ Undesirable☐ Depends☒

Submissive urination

Acceptable☐ Undesirable☒ Depends☐

Excitable urination

Acceptable☐ Undesirable☒ Depends☐

Strong chase behaviours (no grab-bite)

Acceptable☐ Undesirable☐ Depends☒

Barking with retreat

Acceptable☐ Undesirable☒ Depends☐

Barking with approach

Acceptable☐ Undesirable☒ Depends☐

Growling

Acceptable☐ Undesirable☒ Depends☐

Growling with retreat

Acceptable☐ Undesirable☒ Depends☐

Growling with approach

Acceptable☐ Undesirable☒ Depends☐

Snapping

Acceptable☐ Undesirable☒ Depends☐

Biting

Acceptable☐ Undesirable☒ Depends☐

Backing away with low recovery

Acceptable☐ Undesirable☒ Depends☐

Bolting

Acceptable☐ Undesirable☒ Depends☐

Environmental Behaviour

Sound sensitivity

Acceptable☐ Undesirable ☐ Depends☒

Vocal spooking

Acceptable☐ Undesirable ☐ Depends☒

Avoidance

Acceptable☐ Undesirable ☐ Depends☒

Fear/stress behaviours i.e. drooling, shaking, hyperventilating, pacing

Acceptable☐ Undesirable ☒ Depends☐

Backing away with low recovery/Stopping

Acceptable☐ Undesirable ☒ Depends☐

Bolting

Acceptable☐ Undesirable ☒ Depends☐

Copy of CBARQ covering letter given to socialisers

Trial questionnaire for puppy socialising assessments

Dear .....

My name is Hannah Plant. I work within the Quality Assurance department at Hearing dogs for deaf people. I am currently doing a research masters in biology and animal behaviour at Lincoln University. I would like to invite you and your socialising dog to participate in this program and also to try out a new trial questionnaire for Hearing Dogs.

As part of my Masters I will be looking at a sample of 62 Hearing dogs throughout socialising and training. I will be looking closely at how their behaviour in socialising compares to their behaviour in training

Hearing dogs will also be using these questionairres as an initial trial. Information from these questionnaires will be used to hopefully gain a better picture of dogs when they are between 7-9 months old in a socialisers home. This in turn will enable us to recommend any specific needs that the dog may have on entering training, and how we can further help them through their transitional periods. The information you provide may also help us in deciding which recipient the dog would be best suited too.

If the feedback from the questionnaire proves useful, we may in future roll it out to all socialisers and trainers.

I greatly appreciate you taking the time to fill in this questionnaire. Please either return the questionnaire back to me asap at:

C/ O Hannah Plant

Hearing Dogs For Deaf People

Wycombe rd

Saunderton

Bucks

Hp27 9NS

Or scan and Email to me at

Hannah.plant@hearingdogs.org.uk

Alternatively, if it is easier for you, please hand in to your assessor (either myself or Gemma Binstead) at your hearing dogs 8 month assessment.

Please feel free to contact me if you have any further questions

Kind regards

Hannah Plant

Quality Assurance Evaluator

Hearing Dogs For Deaf People

ID Code:

Canine Behavioral Assessment & Research Questionnaire (C–BARQ)

The following questions are designed to allow you to describe how your dog has been behaving in the recent past (i.e. during the last few months). Please try to answer all of the questions. If you have never observed the dog in the situation described, please check the "Not observed/not applicable" box on the right.

SECTION 1: Training difficulty

Some dogs are more obedient and trainable than others. By checking the appropriate boxes, please indicate how trainable or obedient your dog has been in each of the following situations in the recent past:

	Never	Seldom	Sometimes	Usually	Always	Not observed/ applicable
1. When off the leash, returns immediately when called.	☐	☐	☐	☐	☐	☐
2. Obeys the "sit" command immediately.	☐	☐	☐	☐	☐	☐
3. Obeys the "stay" command immediately.	☐	☐	☐	☐	☐	☐
4. Seems to attend/listen closely to everything you say or do.	☐	☐	☐	☐	☐	☐
5. Slow to respond to correction or punishment; 'thick-skinned'.	☐	☐	☐	☐	☐	☐

6. Slow to learn new tricks or tasks.	☐	☐	☐	☐	☐	☐
7. Easily distracted by interesting sights, sounds or smells.	☐	☐	☐	☐	☐	☐
8. Will 'fetch' or attempt to fetch sticks, balls, or objects.	☐	☐	☐	☐	☐	☐

SECTION 2: Aggression

Some dogs display aggressive behavior from time to time. Typical signs of moderate aggression in dogs include barking, growling and baring teeth. More serious aggression generally includes snapping, lunging, biting, or attempting to bite.

By circling or underlining a number on the following 5-point scales (0= No aggression,
4= Serious aggression), please indicate your own dog's recent tendency to display aggressive behavior in each of the following contexts:

9. When verbally corrected or punished (scolded, shouted at, etc) by you or a household member.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

10. When approached directly by an unfamiliar **adult** while being walked/exercised on a leash.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

11. When approached directly by an unfamiliar **child** while being walked/exercised on a leash.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

12. Toward unfamiliar persons approaching the dog while s/he is in your car (at the gas station for example).

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

13. When toys, bones or other objects are taken away by a household member.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

14. When bathed or groomed by a household member.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

15. When an unfamiliar person approaches you or another member of your family at home.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

16. When unfamiliar persons approach you or another member of your family away from your home.

	Moderate aggression:		☐ Not observed
No aggression:	growling/barking—baring teeth	Serious aggression:	
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

17. When approached directly by a household member while s/he (the dog) is eating.

	Moderate aggression:		☐ Not observed
No aggression:	growling/barking—baring teeth	Serious aggression:	
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

18. When mailmen or other delivery workers approach your home.

	Moderate aggression:		☐ Not observed
No aggression:	growling/barking—baring teeth	Serious aggression:	
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

19. When his/her food is taken away by a household member.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

20. When strangers walk past your home while your dog is outside or in the yard.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

21. When an unfamiliar person tries to touch or pet the dog.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

22. When joggers, cyclists, rollerbladers or skateboarders pass your home while your dog is outside or in the yard.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

23. When approached directly by an unfamiliar **male** dog while being walked/exercised on a leash.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

24. When approached directly by an unfamiliar **female** dog while being walked/exercised on a leash.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

25. When stared at directly by a member of the household.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

26. Toward unfamiliar dogs visiting your home.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

27. Toward cats, squirrels or other small animals entering your yard.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

28. Toward unfamiliar persons visiting your home.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

29. When barked, growled, or lunged at by another (unfamiliar) dog.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

30. When stepped over by a member of the household.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

31. When you or a household member retrieves food or objects stolen by the dog.

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

32. Towards another (familiar) dog in your household (leave blank if no other dogs).

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

33. When approached at a favorite resting/sleeping place by another (familiar) household dog (leave blank if no other dogs).

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

34. When approached while eating by another (familiar) household dog (leave blank if no other dogs).

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

35. When approached while playing with/chewing a favorite toy, bone, object, etc., by another (familiar) household dog (leave blank if no other dogs).

No aggression:	Moderate aggression: growling/barking—baring teeth	Serious aggression:	☐ Not observed
No visible signs of aggression	0.....1.....2.....3.....4	Snaps, bites or attempts to bite.	

Are there any other situations in which your dog is sometimes aggressive? If so, please describe briefly:

_____

_____

SECTION 3: Fear and Anxiety

Dogs sometimes show signs of anxiety or fear when exposed to particular sounds, objects, persons or situations. Typical signs of mild to moderate fear include: avoiding eye contact, avoidance of the feared object; crouching or cringing with tail lowered or tucked between the legs; whimpering or whining, freezing, and shaking or trembling. Extreme fear is characterized by exaggerated cowering, and/or vigorous attempts to escape, retreat or hide from the feared object, person or situation.

Using the following 5-point scales (0=No fear, 4=Extreme fear), please indicate your own dog's recent tendency to display fearful behavior in each of the following circumstances:

36. When approached directly by an unfamiliar **adult** while away from your home.

☐ Not observed

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.

37. When approached directly by an unfamiliar **child** while away from your home.

☐ Not observed

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.

38. In response to sudden or loud noises (e.g. vacuum cleaner, car backfire, road drills, objects being dropped, etc.).

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.	

39. When unfamiliar persons visit your home.

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.	

40. When an unfamiliar person tries to touch or pet the dog.

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.	

41. In heavy traffic

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible		cowers; retreats or	

signs of fear 0.....1.....2.....3.....4 hides, etc.

42. In response to strange or unfamiliar objects on or near the sidewalk (e.g. plastic trash bags, leaves, litter, flags flapping, etc.

☐ Not observed

No fear/anxiety: Mild—Moderate fear/anxiety Extreme fear:

No visible cowers; retreats or hides, etc.

signs of fear 0.....1.....2.....3.....4

43. When examined/treated by a veterinarian.

☐ Not observed

No fear/anxiety: Mild—Moderate fear/anxiety Extreme fear:

No visible cowers; retreats or hides, etc.

signs of fear 0.....1.....2.....3.....4

44. During thunderstorms, firework displays, or similar events.

☐ Not observed

No fear/anxiety: Mild—Moderate fear/anxiety Extreme fear:

No visible cowers; retreats or hides, etc.

signs of fear 0.....1.....2.....3.....4

45. When approached directly by an unfamiliar dog of the same or larger size.

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.	

46. When approached directly by an unfamiliar dog of a smaller size.

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.	

47. When first exposed to unfamiliar situations (e.g. first car trip, first time in elevator, first visit to veterinarian, etc.)

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.	

48. In response to wind or wind-blown objects.

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible		cowers; retreats or	

signs of fear 0.....1.....2.....3.....4 hides, etc.

49. When having nails clipped by a household member.

☐ Not observed

No fear/anxiety: Mild—Moderate fear/anxiety Extreme fear:

No visible signs of fear 0.....1.....2.....3.....4 cowers; retreats or hides, etc.

50. When groomed or bathed by a household member.

☐ Not observed

No fear/anxiety: Mild—Moderate fear/anxiety Extreme fear:

No visible signs of fear 0.....1.....2.....3.....4 cowers; retreats or hides, etc.

51. When having his/her feet toweled by a member of the household.

☐ Not observed

No fear/anxiety: Mild—Moderate fear/anxiety Extreme fear:

No visible signs of fear 0.....1.....2.....3.....4 cowers; retreats or hides, etc.

52. When unfamiliar dogs visit your home.

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.	

53. When barked, growled, or lunged at by an unfamiliar dog.

No fear/anxiety:	Mild—Moderate fear/anxiety	Extreme fear:	☐ Not observed
No visible signs of fear	0.....1.....2.....3.....4	cowers; retreats or hides, etc.	

SECTION 4: Separation-related behavior.

Some dogs show signs of anxiety or abnormal behavior when left alone, even for relatively short periods of time. Thinking back over the recent past, how often has your dog shown each of the following signs of separation-related behavior when left, or about to be left, on its own (check appropriate boxes):

	Never	Seldom	Sometimes	Usually	Always	Not observed/ not applicable
54. Shaking, shivering or trembling.	☐	☐	☐	☐	☐	☐
55. Excessive salivation.	☐	☐	☐	☐	☐	☐
56. Restlessness/agitation/pacing.	☐	☐	☐	☐	☐	☐
57. Whining.	☐	☐	☐	☐	☐	☐
58. Barking.	☐	☐	☐	☐	☐	☐
59. Howling.	☐	☐	☐	☐	☐	☐
60. Chewing/scratching at doors, floor, windows, curtains, etc.	☐	☐	☐	☐	☐	☐
61. Loss of appetite.	☐	☐	☐	☐	☐	☐

Are there any other situations in which your dog is fearful or anxious? If so, please describe:

_____

_____

SECTION 5: Excitability

Some dogs show relatively little reaction to sudden or potentially exciting events and disturbances in their environment, while others become highly excited at the slightest novelty. Signs of mild to moderate excitability include increased alertness, movement toward the source of novelty, and brief episodes of barking. Extreme excitability is characterized by a general tendency to over-react. The excitable dog barks or yelps hysterically at the slightest disturbance, rushes towards and around any source of excitement, and is difficult to calm down.

Using the following 5-point scales (0=Calm, 4=Extremely excitable), please indicate your own dog's recent tendency to become excitable in each of the following circumstances:

62. When you or other members of the household come home after a brief absence.

Calm: little or no special reaction	Mild—Moderate excitability	Extremely excitable: over-reacts, hard to calm down.	☐ Not observed
	0.....1.....2.....3.....4		

63. When playing with you or other members of your household.

Calm: little or no special reaction

Mild—Moderate excitability

Extremely excitable: over-reacts, hard to calm down.

0.....1.....2.....3.....4

☐ Not observed

64. When doorbell rings.

Calm: little or no special reaction

Mild—Moderate excitability

Extremely excitable: over-reacts, hard to calm down.

0.....1.....2.....3.....4

☐ Not observed

65. Just before being taken for a walk.

Calm: little or no special reaction

Mild—Moderate excitability

Extremely excitable: over-reacts, hard to calm down.

0.....1.....2.....3.....4

☐ Not observed

66. Just before being taken on a car trip.

Calm: little or no special reaction

Mild—Moderate excitability

Extremely excitable: over-reacts, hard to calm down.

0.....1.....2.....3.....4

☐ Not observed

67. When visitors arrive at your home.

Calm: little or no special reaction	Mild—Moderate excitability	Extremely excitable: over-reacts, hard to calm down.	☐ Not observed
	0.....1.....2.....3.....4		

Are there any other situations in which your dog sometimes becomes over-excited? If so, please describe briefly:

_____

SECTION 6: Attachment and Attention-seeking.

Most dogs are strongly attached to their people, and some demand a great deal of attention and affection from them. Thinking back over the recent past, how often has your dog shown each of the following signs of attachment or attention-seeking.

	Never	Seldom	Sometimes	Usually	Always	Not observed/ not applicable
68. Displays a strong attachment for one particular member of the household.	☐	☐	☐	☐	☐	☐
69. Tends to follow you (or other members of household) about the house, from room to room.	☐	☐	☐	☐	☐	☐

70. Tends to sit close to, or in contact with, you (or others) when you are sitting down.	☐	☐	☐	☐	☐	☐
71. Tends to nudge, nuzzle or paw you (or others) for attention when you are sitting down.	☐	☐	☐	☐	☐	☐
72. Becomes agitated (whines, jumps up, tries to intervene) when you (or others) show affection for another person.	☐	☐	☐	☐	☐	☐
73. Becomes agitated (whines, jumps up, tries to intervene) when you show affection for another dog or animal.	☐	☐	☐	☐	☐	☐

SECTION 7: Miscellaneous

Dogs display a wide range of miscellaneous behavior problems in addition to those already covered by this questionnaire. Thinking back over the recent past, please indicate how often your dog has shown any of the following behaviors:

	Never	Seldom	Sometimes	Usually	Always	Not observed/ not applicable
74. Chases or would chase cats given the opportunity.	☐	☐	☐	☐	☐	☐
75. Chases or would chase birds given the opportunity.	☐	☐	☐	☐	☐	☐
76. Chases or would chase squirrels, rabbits and other small animals given the opportunity.	☐	☐	☐	☐	☐	☐

77. Escapes or would escape from home or yard given the chance.	☐	☐	☐	☐	☐	☐
78. Rolls in animal droppings or other ‘smelly’ substances.	☐	☐	☐	☐	☐	☐
79. Eats own or other animals’ droppings or feces.	☐	☐	☐	☐	☐	☐
80. Chews inappropriate objects.	☐	☐	☐	☐	☐	☐
81. ‘Mounts’ objects, furniture, or people.	☐	☐	☐	☐	☐	☐
82. Begs persistently for food when people are eating.	☐	☐	☐	☐	☐	☐
83. Steals food.	☐	☐	☐	☐	☐	☐
84. Nervous or frightened on stairs.	☐	☐	☐	☐	☐	☐
85. Pulls excessively hard when on the leash.	☐	☐	☐	☐	☐	☐
86. Urinates against objects/ furnishings in your home.	☐	☐	☐	☐	☐	☐

	Never	Seldom	Sometimes	Usually	Always	not applicable
87. Urinates when approached, petted, handled or picked up.	☐	☐	☐	☐	☐	☐
88. Urinates when left alone at night, or during the daytime.	☐	☐	☐	☐	☐	☐
89. Defecates when left alone at night, or during the daytime.	☐	☐	☐	☐	☐	☐
90. Hyperactive, restless, has trouble settling down.	☐	☐	☐	☐	☐	☐
91. Playful, puppyish, boisterous.	☐	☐	☐	☐	☐	☐
92. Active, energetic, always on the go.	☐	☐	☐	☐	☐	☐
93. Stares intently at nothing visible.	☐	☐	☐	☐	☐	☐
94. Snaps at (invisible) flies.	☐	☐	☐	☐	☐	☐
95. Chases own tail/hind end.	☐	☐	☐	☐	☐	☐
96. Chases/follows shadows, light spots, etc.	☐	☐	☐	☐	☐	☐
97. Barks persistently when alarmed or excited.	☐	☐	☐	☐	☐	☐

98. Licks him/herself excessively.	☐	☐	☐	☐	☐	☐
99. Licks people or objects excessively.	☐	☐	☐	☐	☐	☐
100. Displays other bizarre, strange, or repetitive behavior(s) *	☐	☐	☐	☐	☐	☐

* Please describe: _____

C-BARQ$^{(100)}$ scoring method

The C-*BARQ* provides a set of quantitative scores for the following fourteen different subscales or categories of behavior:

1. ***Stranger-directed aggression*****:** Dog shows threatening or aggressive responses to strangers approaching or invading the dog's or the owner's personal space, territory, or home range.
2. ***Owner-directed aggression*****:** Dog shows threatening or aggressive responses to the owner or other members of the household when challenged, manhandled, stared at, stepped over, or when approached while in possession of food or objects.
3. ***Dog-directed aggression:*** Dog shows threatening or aggressive responses when approached directly by unfamiliar dogs.
4. ***Dog rivalry*****:** Dog shows aggressive or threatening responses to other familiar dogs in the same household.
5. ***Stranger-directed fear*****:** Dog shows fearful or wary responses when approached directly by strangers.
6. ***Nonsocial fear*****:** Dog shows fearful or wary responses to sudden or loud noises (e.g. thunder), traffic, and unfamiliar objects and situations.
7. ***Dog-directed fear*****:** Dog shows fearful or wary responses when approached directly by unfamiliar dogs.
8. ***Touch sensitivity*****:** Dog shows fearful or wary responses to potentially painful or uncomfortable procedures, including bathing, grooming, nail-clipping, and veterinary examinations.
9. ***Separation-related behavior*****:** Dog vocalizes and/or is destructive when separated from the owner, often accompanied or preceded by behavioral and autonomic signs of anxiety including restlessness, loss of appetite, trembling, and excessive salivation.
10. ***Attachment and attention-seeking*****:** Dog maintains close proximity to the owner or other members of the household, solicits affection or attention, and displays agitation when the owner gives attention to third parties.
11. ***Trainability*****:** Dog shows a willingness to attend to the owner and obey simple commands. Dog is not easily distracted, tends to be a fast learner, responds positively to correction, and will fetch or retrieve objects.
12. ***Chasing*****:** Dog chases cats, birds, and/or other small animals, given the opportunity.
13. ***Excitability*****:** Dog displays strong reaction to potentially exciting or arousing events, such as going for walks or car trips, doorbells, arrival of visitors, and the owner arriving home; has difficulty calming down after such events.
14. ***Energy level*****:** Dog is energetic, "always on the go", and/or playful.

In addition, the C-BARQ provides useful information on the occurrence of a further 22 miscellaneous behavior problems ranging from coprophagia to stereotypic spinning/tail-chasing.

Each subscale is represented by a number of 5-point scales (questions). Some are graduated scales that measure severity of particular behaviors (e.g. aggression, fear, excitability) and are numbered from 0–4 in the questionnaire. The remainder are frequency scales which should be scored as: Never

= 0, Seldom = 1, Sometimes = 2, Usually = 3 and Always = 4, **except for items 5, 6 & 7 in Section 1. FOR THESE SCALES ONLY, reverse the scores to: Never = 4, Seldom = 3, etc.** To calculate behavior subscale scores, use the following formulae:

"Stranger-directed aggression" score = (questionnaire items 10 + 11 + 12 + 15 + 16 + 18 + 20 + 21 + 22 + 28)/10.

"Owner-directed aggression" score = (items 9 + 13 + 14 + 17 + 19 + 25 + 30 + 31)/8.

"Dog-directed aggression" = (items 23 + 24 + 26 + 29)/4

"Dog-directed fear" = (items 45 + 46 + 52 + 53)/4.

"Dog rivalry"(familiar dog aggression) score = (items 32 + 33 + 34 + 35)/4

"Trainability" score = (items 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8)/8—**remember to reverse scoring order for items 5, 6 & 7 (see above).**

"Chasing" score = (items 27 + 74 + 75 + 76)/4

"Stranger-directed fear" score = (items 36 + 37 + 39 + 40)/4

"Nonsocial fear" score = (items 38 + 41 + 42 + 44 + 47 + 48)/6

"Separation-related problems" score = (items 54 + 55 + 56 + 57 + 58 + 59 + 60 + 61)/8

"Touch sensitivity" score = (items 43 + 49 + 50 + 51)/4

"Excitability" score = (items 62 + 63 + 64 + 65 + 66 + 67)/6

"Attachment/attention-seeking" score = (items 68 + 69 + 70 + 71 + 72 + 73)/6

"Energy" score = (items 91 + 92)/2

Items 1–76 & 91–92 cannot be removed from the questionnaire without potentially reducing the reliability and/or validity of one or other of the behavior subscales. Other "Miscellaneous" items are optional, and can be removed from the questionnaire as desired. If retained, they should be scored individually, 0–4.

Missing values: Owners may be unable to answer some of the C-BARQ questions for a variety of reasons. These "Not observed/Not applicable" responses should be recorded as missing values and the subscale scores calculated as the average of the remaining completed item scores. If more than 25% of the items in a subscale are missing values, the factor/subscale score should be recorded as a missing value.

NB: This version of the C-*BARQ* has been modified since Hsu & Serpell (2003) to improve the reliability of some existing factors, and to include new "Dog rivalry (familiar dog aggression)" and "Energy" factors. The subscales "Dog rivalry", "Chasing", "Touch sensitivity", "Trainability", "Energy" and "Excitability" have not been formally validated, although they have been shown to have predictive validity in long-term studies of working guide dogs (Duffy & Serpell, 2012).

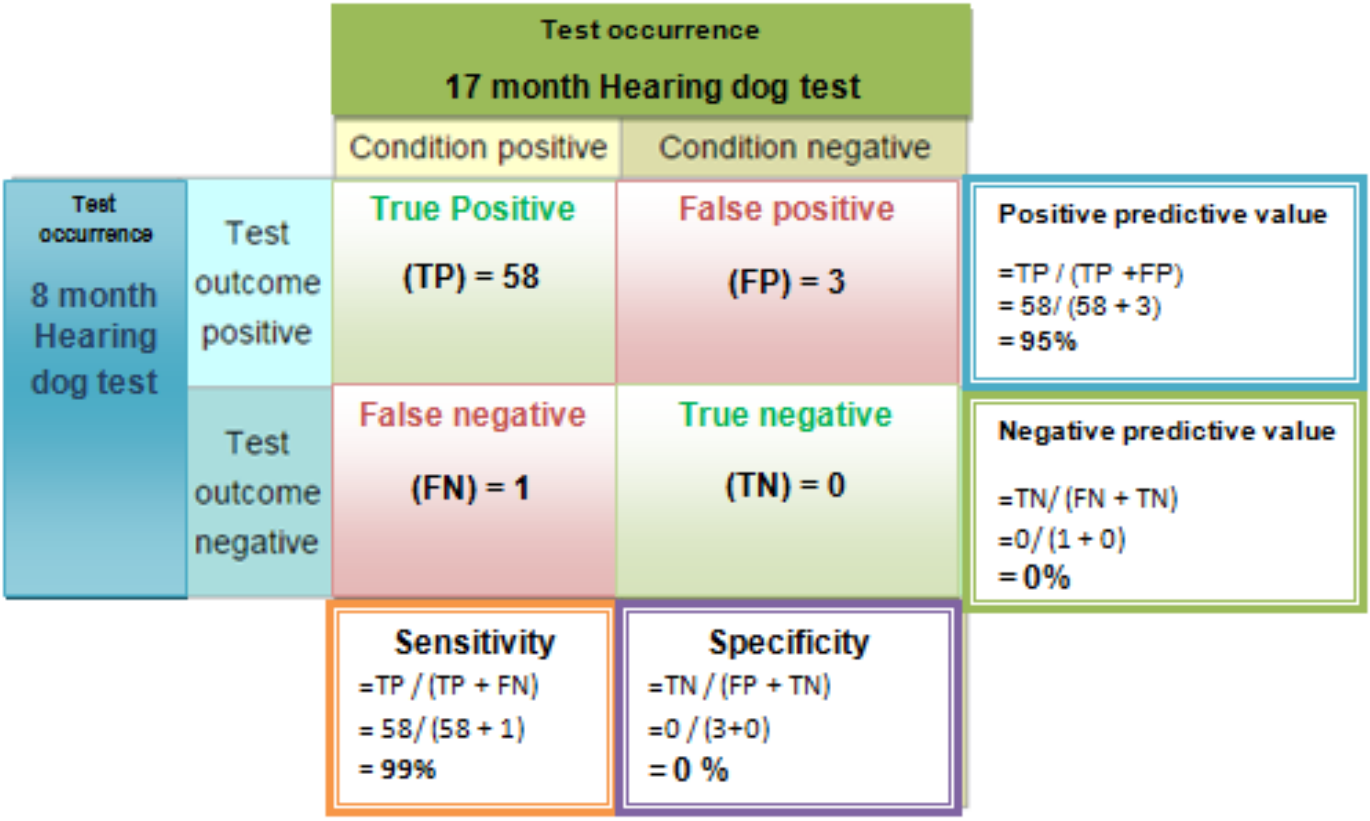
References:

Hsu, Y. and Serpell, J.A. 2003. Development and validation of a questionnaire for measuring behavior and temperament traits in pet dogs. *Journal of the American Veterinary Medical Association*, 223: 1293-1300.

Duffy, D.L. and Serpell, J.A. 2012. Predictive validity of a method for evaluating temperament in young guide and service dogs. *Applied Animal Behavior Science*, 138: 99-109.

Sensitivity and Specificity Tables

1) A diagnostic test to examine the Hearing dog test for social behaviour with adults



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **59**

Total that did not pass their 17 month HD social adult test = **3**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 3 / (3 + 0) = **100%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 1 / (58 + 0) = **2%**

Power = sensitivity = 1− β

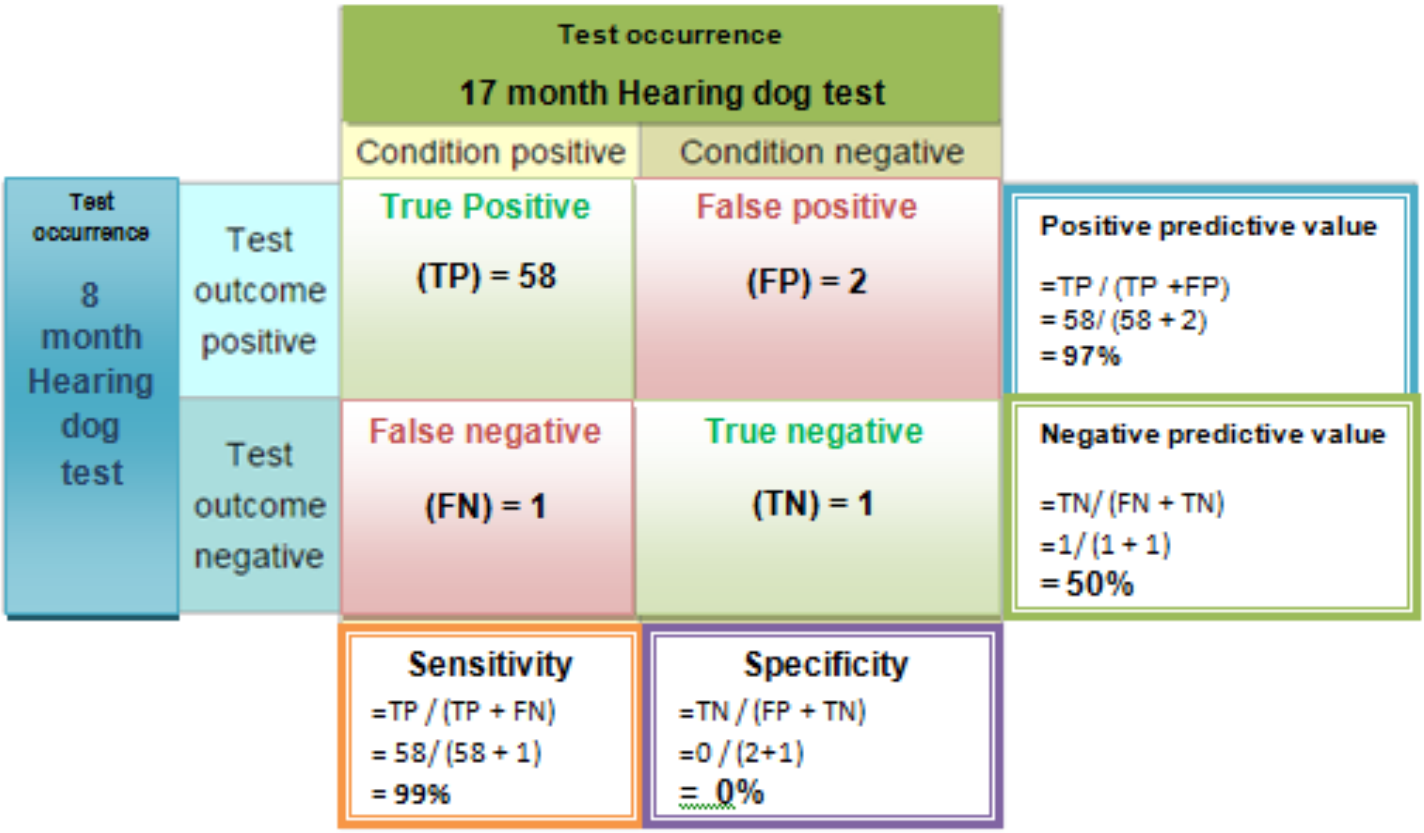
Likelihood ratio positive = sensitivity / (1- specificity) = 0.99 / (1 – 0) = **0.99**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.99) / 0 = **not calculable**

Positive post test probability= 95%

Negative post test probability = 0%

2) A diagnostic test to examine the Hearing dog test for social behaviour with Children



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **59**

Total that did not pass their 17 month HD social adult test = **3**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 2 / (2+ 1) = 67 **%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 1 / (58 + 1) =2**%**

Power = sensitivity = 1− β

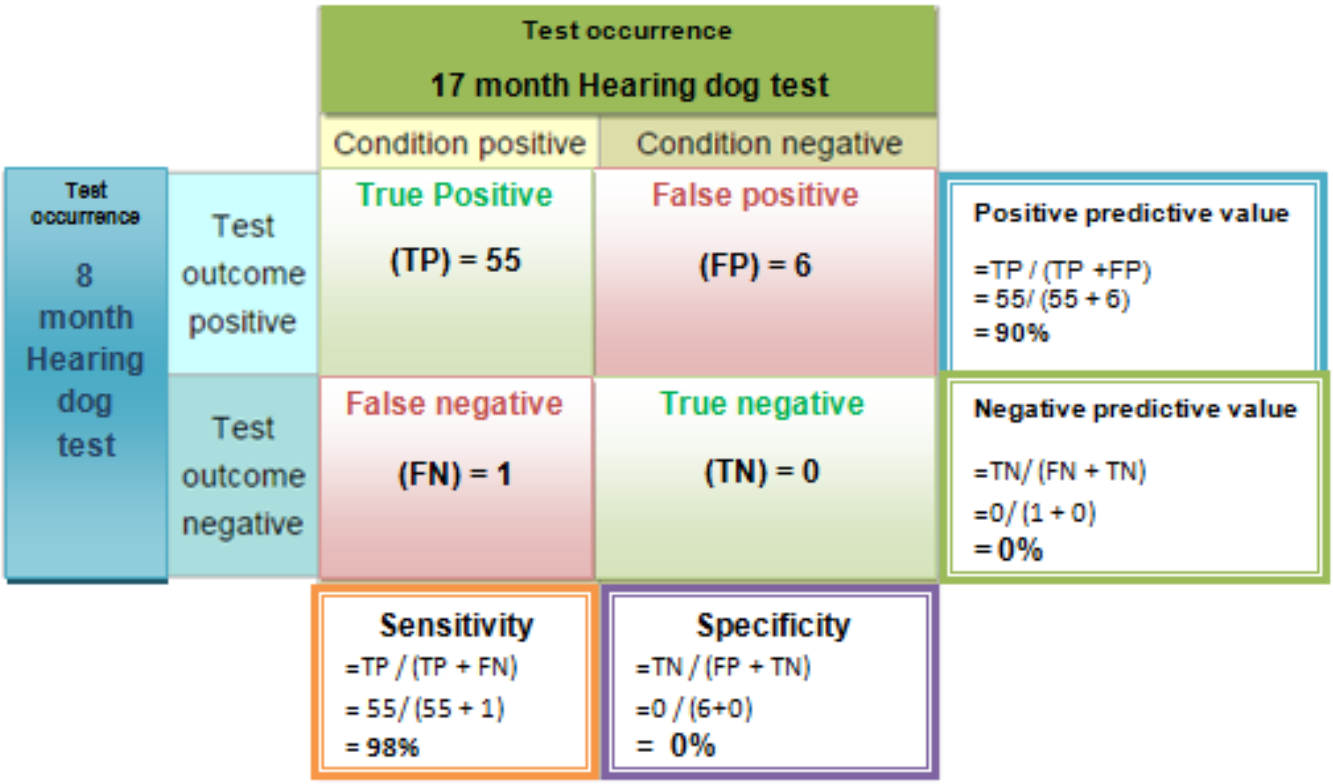
Likelihood ratio positive = sensitivity / (1- specificity) = 0.99 / (1 – 0) = **0.99**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.99) / 0 = **not calculable**

Positive post test probability = 97%

Negative post test probability = 50%

3. A diagnostic test to examine the Hearing dog test for social behaviour with dogs



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **56**

Total that did not pass their 17 month HD social adult test = **6**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 6 / (6+ 0) = 100**%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 1 / (55 + 1) = 2**%**

Power = sensitivity = 1− β

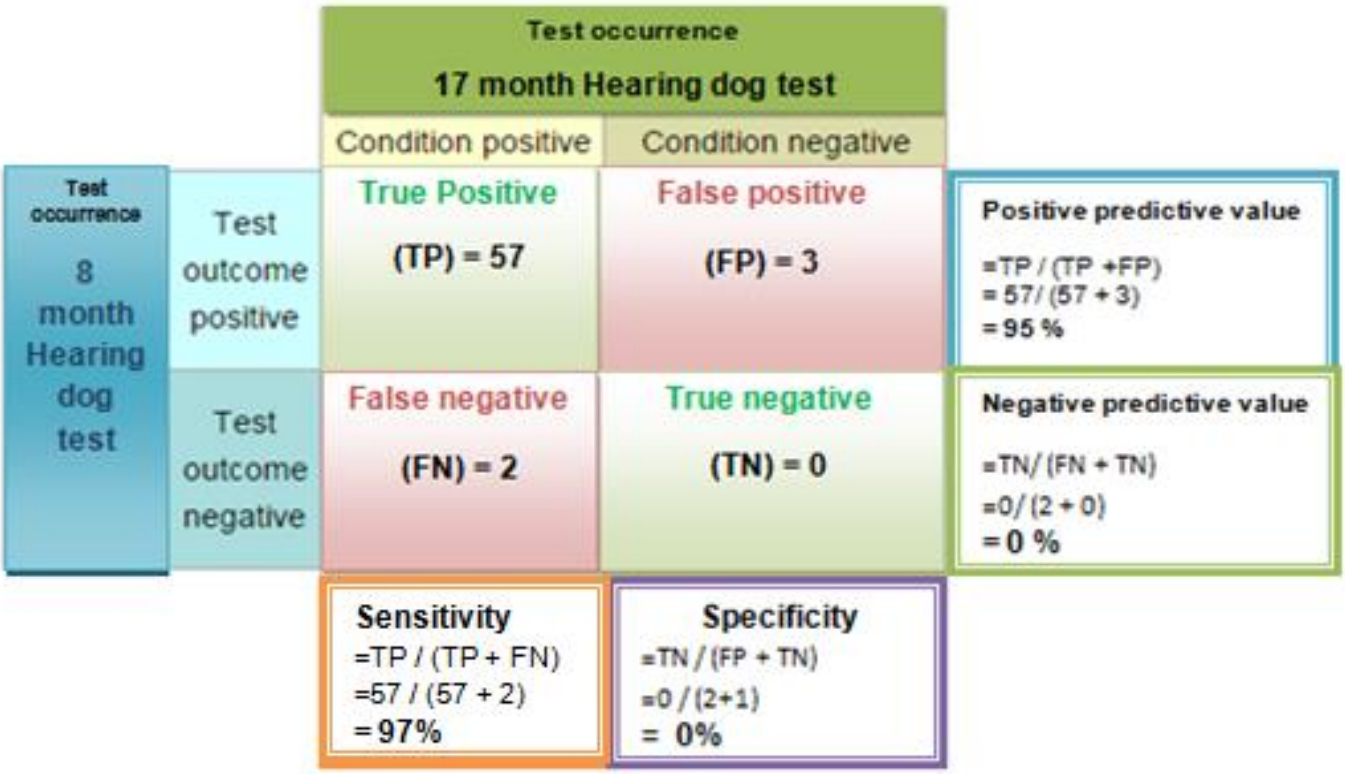
Likelihood ratio positive = sensitivity / (1- specificity) = 0.98 / (1 – 0) = **0.98**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.98) / 0 = **not calculable**

Positive Post test probability = 89%

Negative post test probability = 0%

4. A diagnostic test to examine the Hearing dog test for Environmental confidence



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **59**

Total that did not pass their 17 month HD social adult test = **3**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 3 / (3+ 0) = 100**%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 2 / (57 + 2) = 3**%**

Power = sensitivity = 1− β

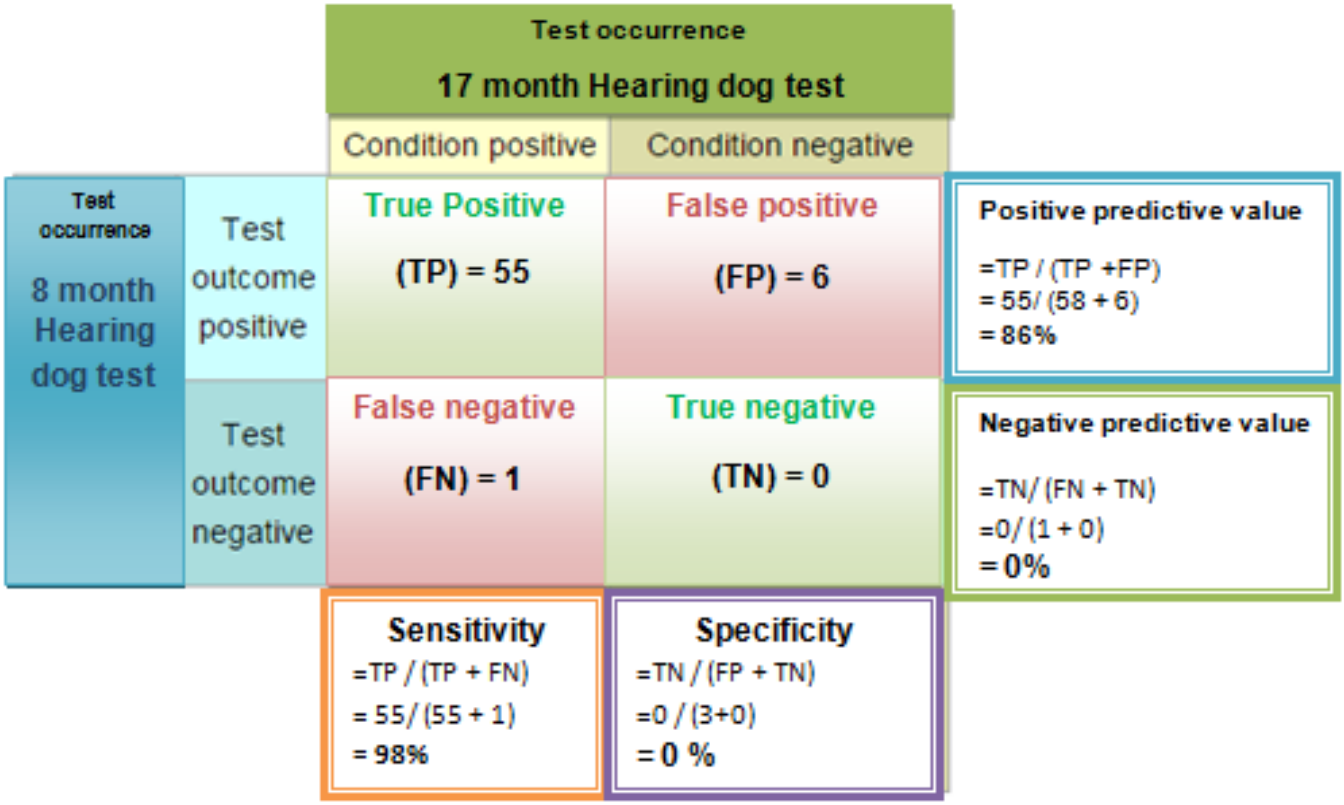
Likelihood ratio positive = sensitivity / (1- sensitivity) = 0.97 / (1 – 0) = **0.9**7

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.97) / 0 = **not calculable**

Post test probability = 94%

Negative post test probability = 0%

5. A diagnostic test to examine the Hearing dog test for Recovery rate



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **56**

Total that did not pass their 17 month HD social adult test = **6**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 6 / (6+ 0) = 100**%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 1 / (55 + 1) = 2**%**

Power = sensitivity = 1− β

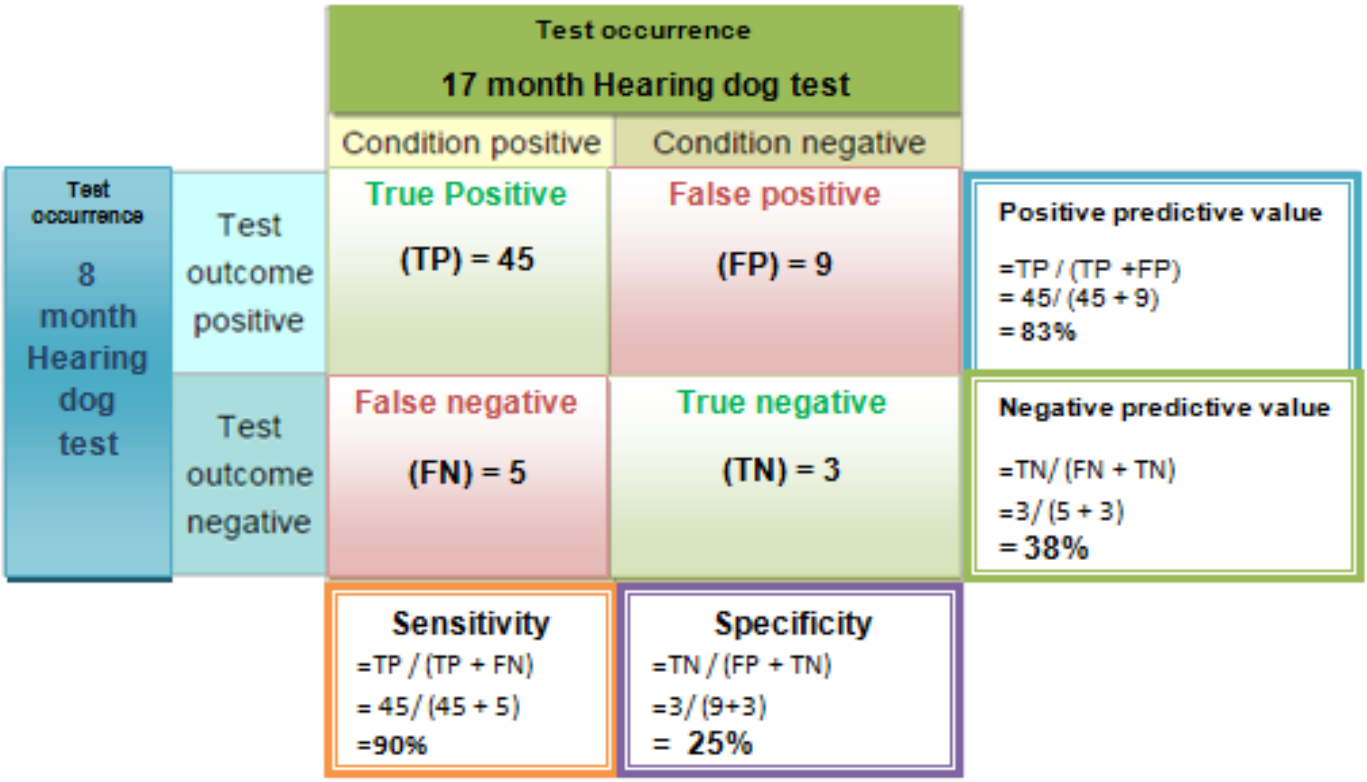
Likelihood ratio positive = sensitivity / (1- specificity) = 0.98 / (1 – 0) = **0.9**8

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.98) / 0 = **not calculable**

Positive Post test probability = 86%

Negative post test probability = 0%

6. A diagnostic test to examine the Hearing dog test for adaptability



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **50**

Total that did not pass their 17 month HD social adult test = **12**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 9 / (9+ 3) = 75**%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 5 / (45 + 5) = 10**%**

Power = sensitivity = 1− β

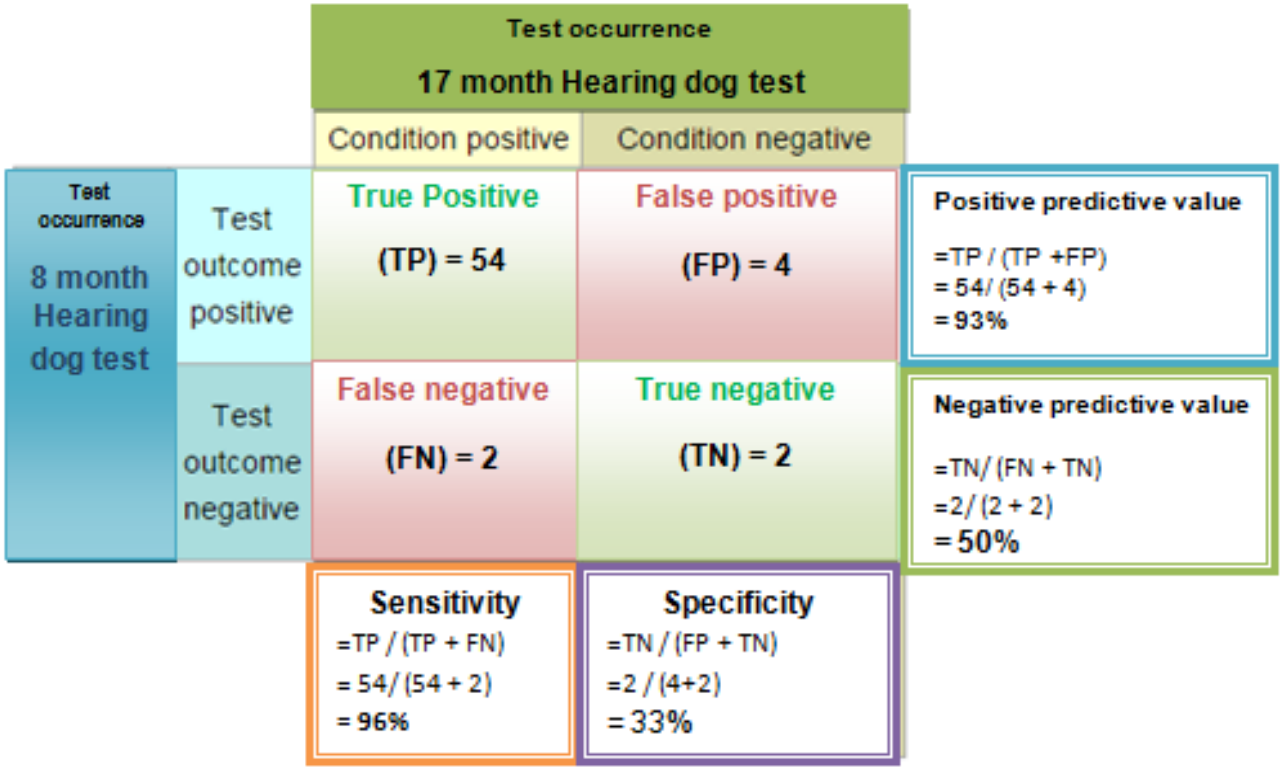
Likelihood ratio positive = sensitivity / (1- specificity) = 0.90 / (1 – 0.25) =**- 1.2**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.90) / 0.25 = **0.4**

Positive post test probability = 85%

Negative post test probability = 20%

7. A diagnostic test to examine the Hearing dog test for Vocal reactivity



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **56**

Total that did not pass their 17 month HD social adult test = **6**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 4 / (4+ 2) = 67**%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 2 / (54 + 2) = 4**%**

Power = sensitivity = 1− β

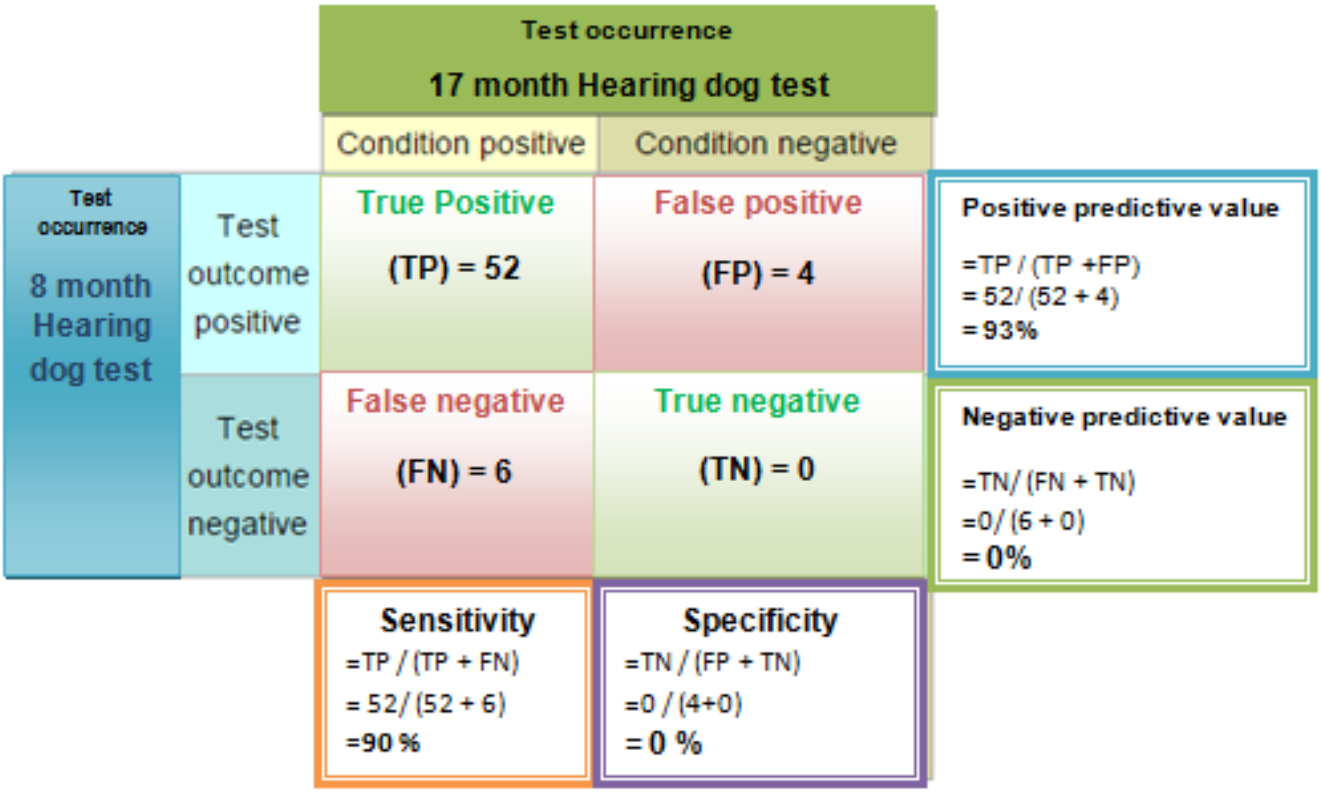
Likelihood ratio positive = sensitivity / (1- specificity) = 0.96 / (1 – 0.33) = **1.43**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.96) / 0.33 = **0.12**

Positive post test probability= 95%

Negative post test probability = 11%

8. A diagnostic test to examine the Hearing dog test for frustration



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **52**

Total that did not pass their 17 month HD social adult test = **4**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 4 / (4+ 0) =100 **%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 6 / (52 + 6) = 10**%**

Power = sensitivity = 1− β

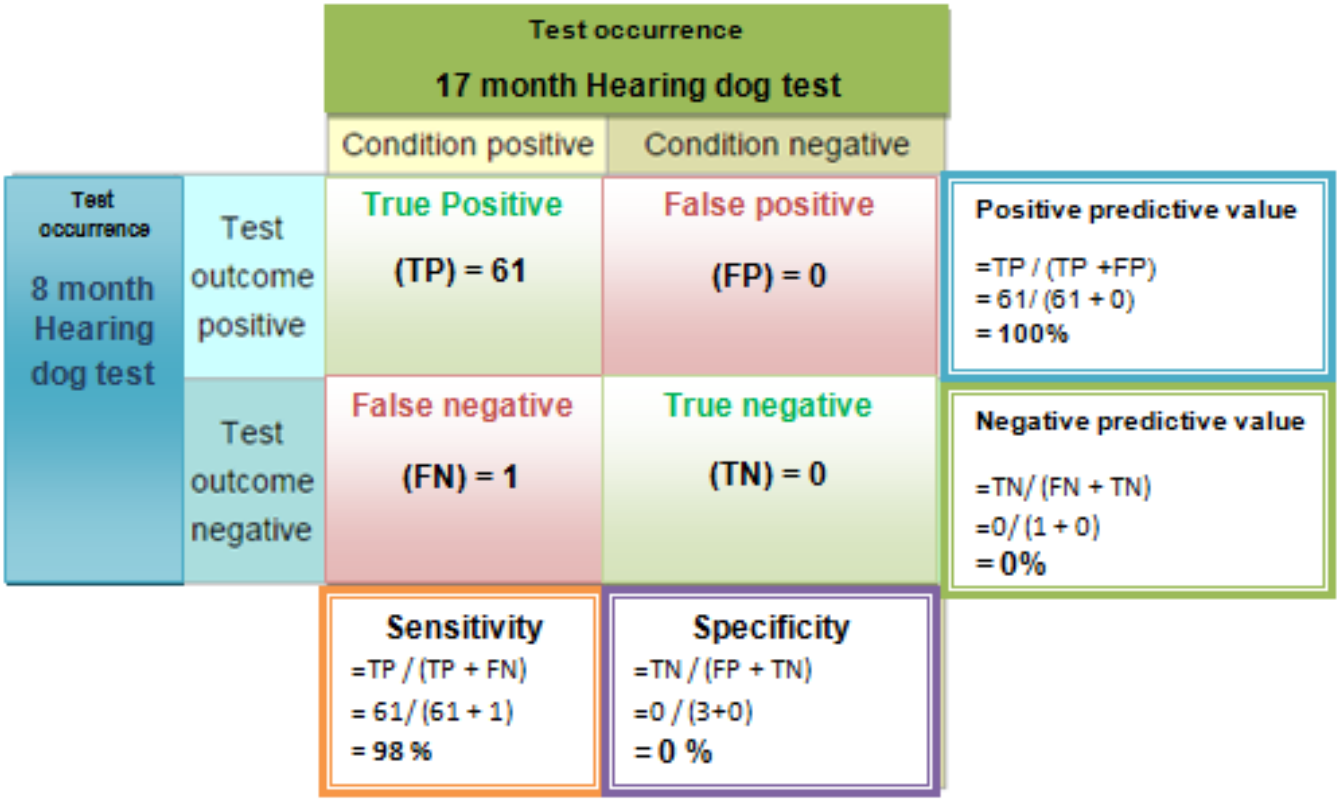
Likelihood ratio positive = sensitivity / (1- specificity) = 0.90 / (1 – 0) = **0.90**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.90) / **not calculable**

Positive post test probability = 92%

Negative post test probability = 0%

9. A diagnostic test to examine the Hearing dog test for motivation



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **62**

Total that did not pass their 17 month HD social adult test = **0**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 0 / (0+ 0) = not calculable

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 1 / (61 + 1) = 2**%**

Power = sensitivity = 1− β

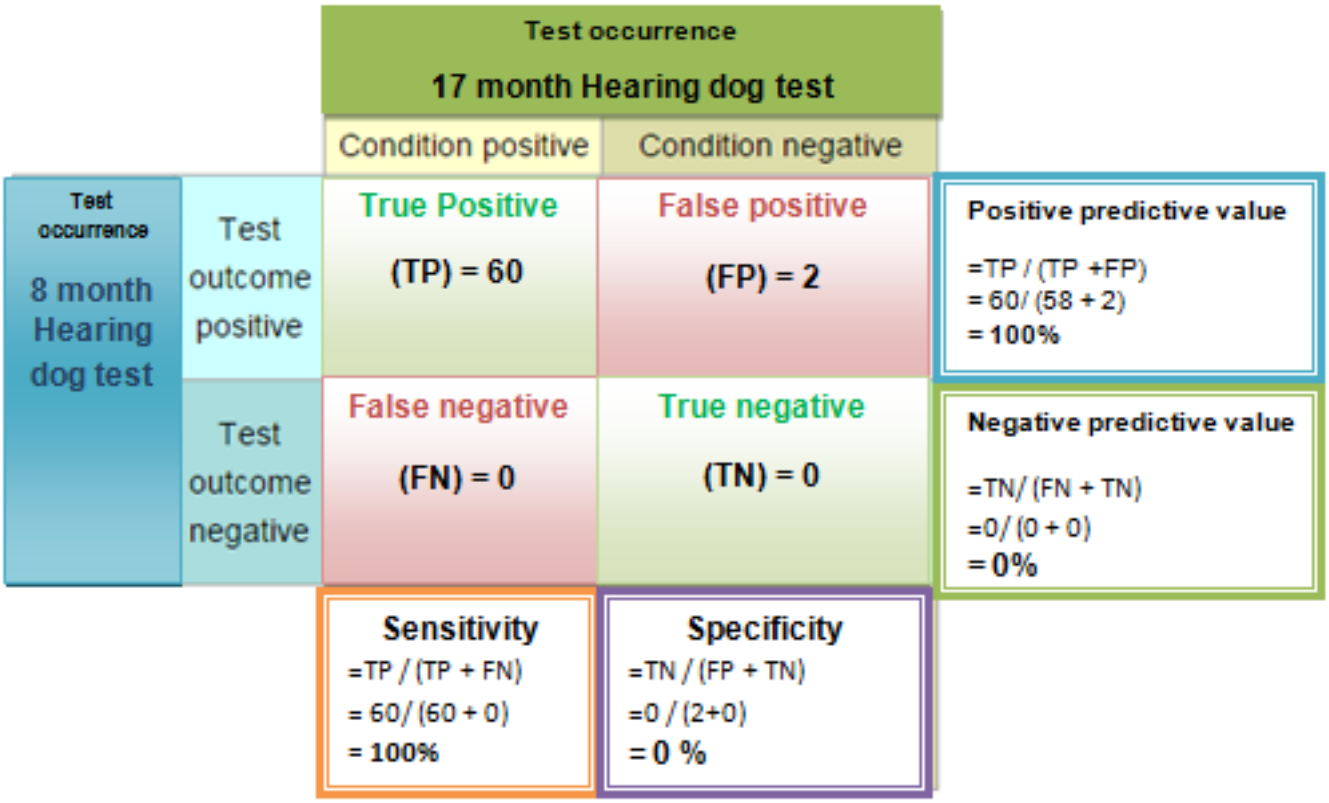
Likelihood ratio positive = sensitivity / (1- specificity) = 0.98 / (1 – 0) = **0.98**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.98) / 0 = **not calculable**

Positive post test probability = 100%

Negative post test probability = 0%

10. A diagnostic test to examine the Hearing dog test for trainability



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **60**

Total that did not pass their 17 month HD social adult test = **2**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 2 / (2+ 0) = 100**%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 0 / (60 + 0) = not calculable

Power = sensitivity = 1− β

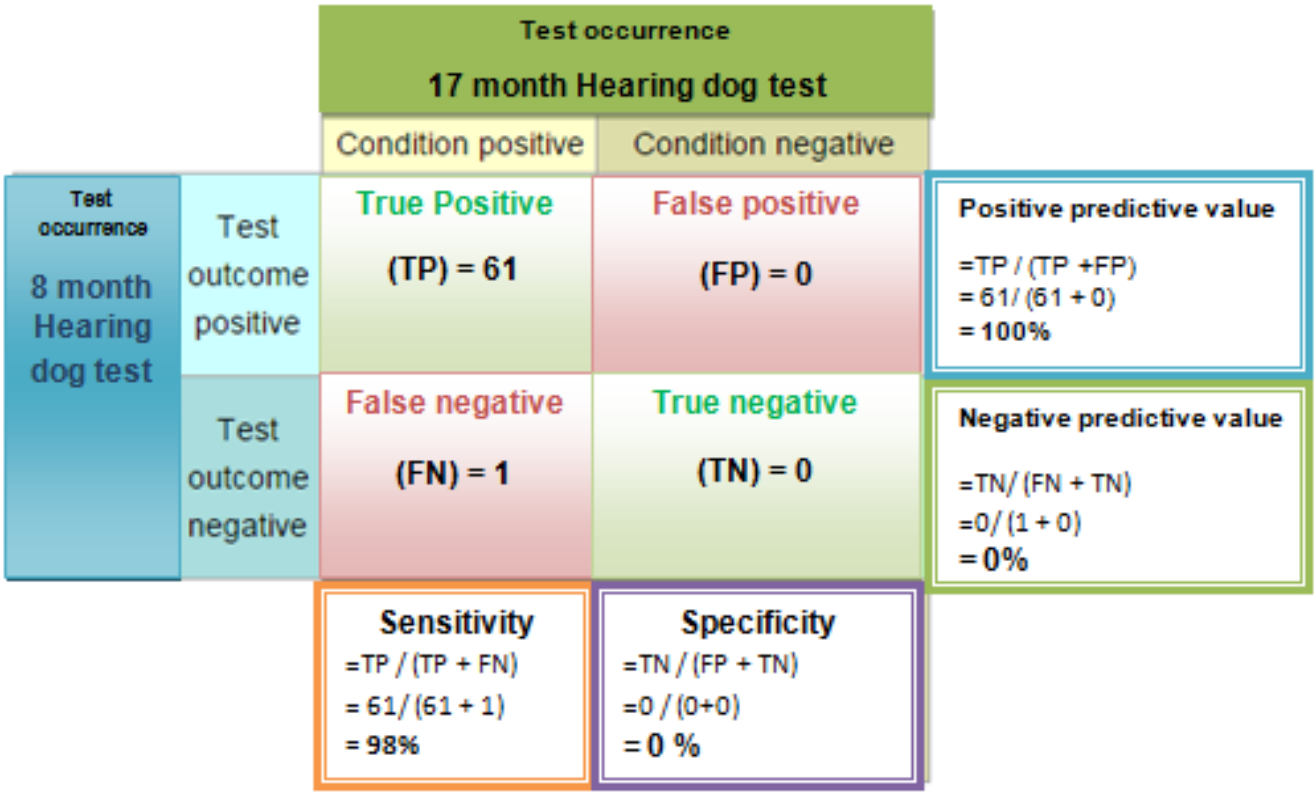
Likelihood ratio positive = sensitivity / (1- specificity) = 1/ (1 – 0) = **1**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 1) / 0 = **not calculable**

Positive post test probability = 100%

Negative post test probability= 0%

11. A diagnostic test to examine the Hearing dog test for chase



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **62**

Total that did not pass their 17 month HD social adult test = **0**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 0 / (0+ 0) = not calculable

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 1/ (61 + 1) = 2**%**

Power = sensitivity = 1− β

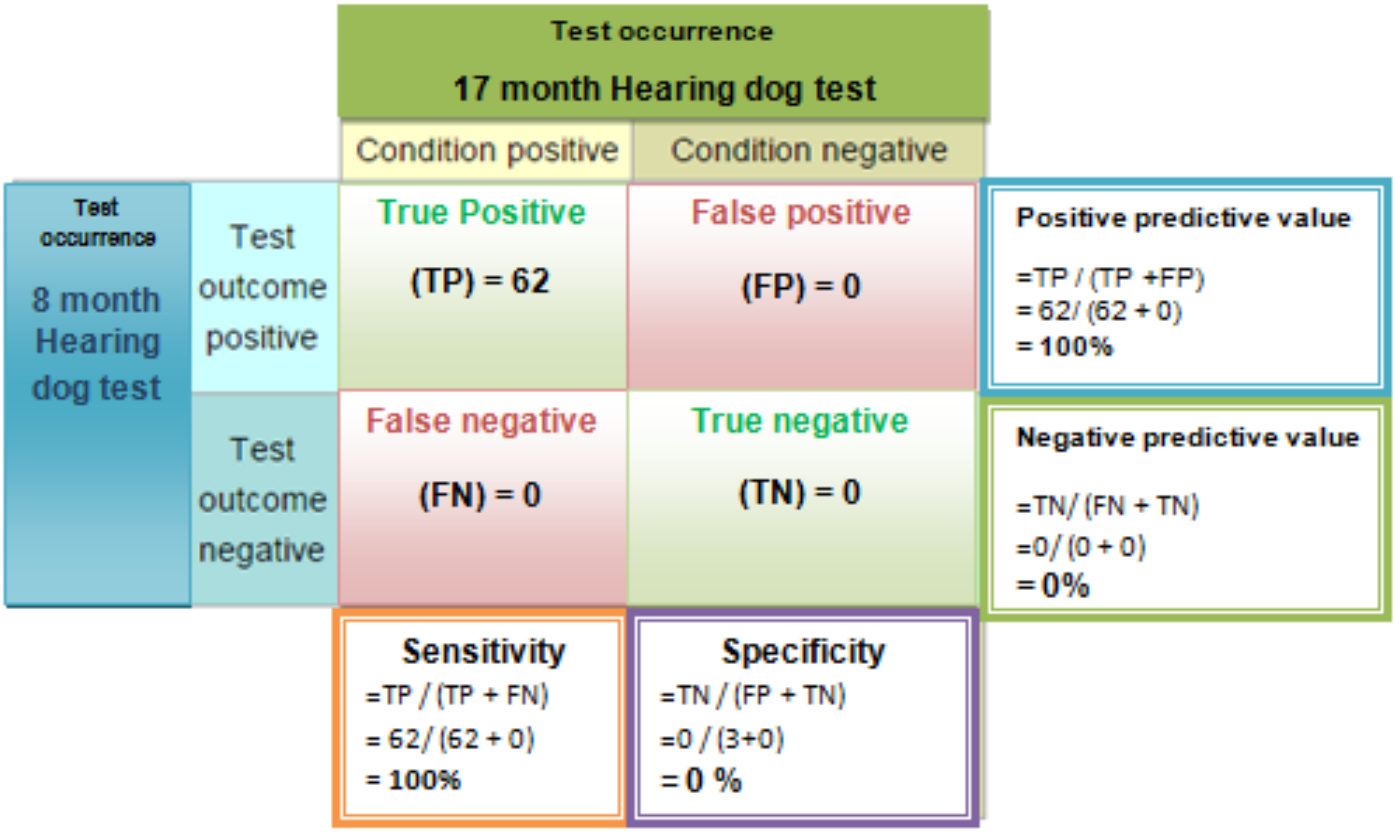
Likelihood ratio positive = sensitivity / (1- specificity) = 0.98/ (1 – 0.98) = **0.49**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.98) / 0 = **not calculable**

Positive post test probability = 100%

Negative post test probability = 0%

12. A diagnostic test to examine the Hearing dog test for hunt



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **62**

Total that did not pass their 17 month HD social adult test = **0**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 0 / (0+ 0) = not calculable

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 0/ (62 + 0) = not calculable

Power = sensitivity = 1− β

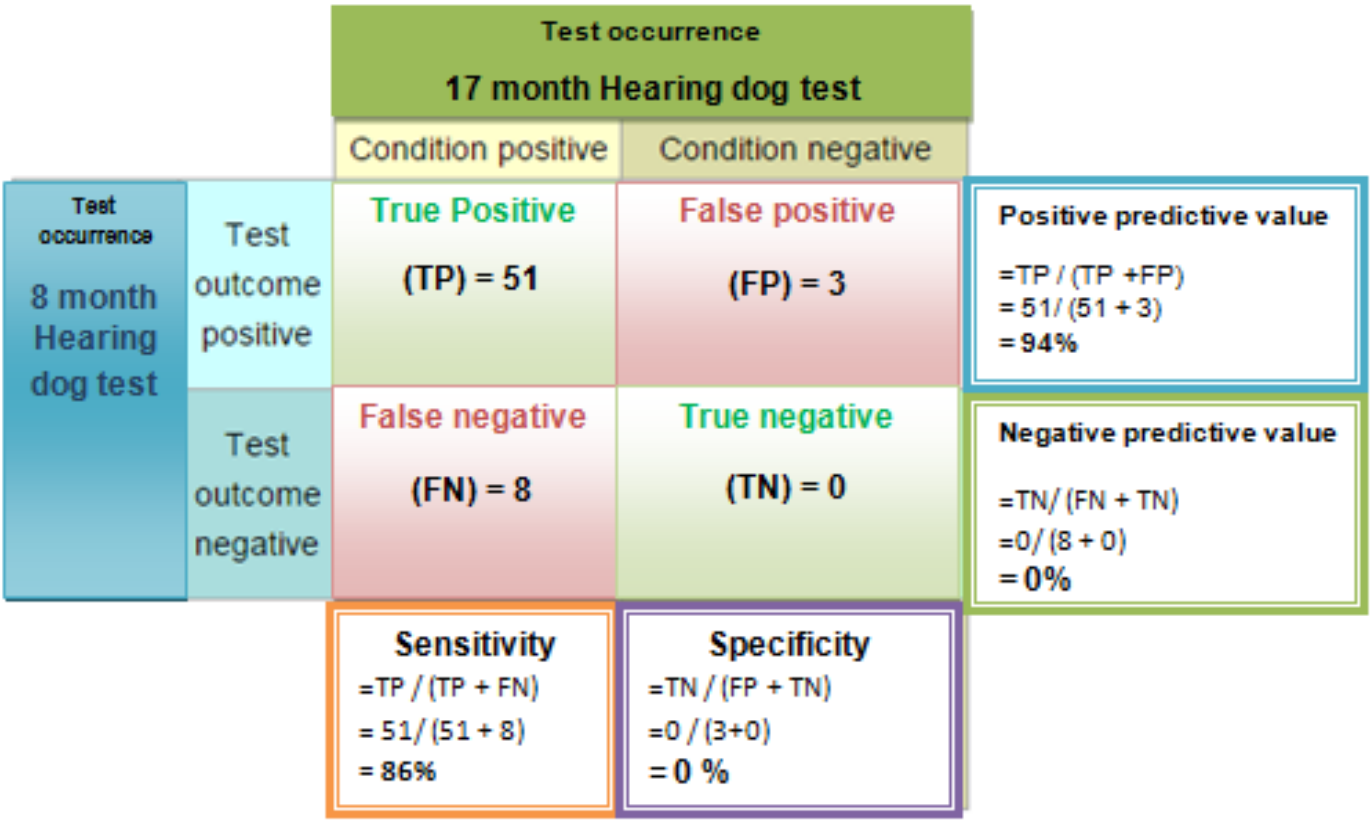
Likelihood ratio positive = sensitivity / (1- specificity) = 1/ (1 – 0) = **1**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 1) / 0 = **not calculable**

Positive post test probability = 100%

Negative post test probability = O%

13) A diagnostic test to examine the Hearing dog test for distraction



Related calculations

Total dogs that were tested =**62**

Total that passed their 17 month HD social adult test = **59**

Total that did not pass their 17 month HD social adult test = **3**

False positive rate (α) = type I error= 1- specificity =FP / (FP + TN) = 3 / (3+ 0) = 100**%**

False negative rate (β) = type II error = 1 – sensitivity = FN / (TP + FN) = 8/ (51 + 8) = 14**%**

Power = sensitivity = 1− β

Likelihood ratio positive = sensitivity / (1- sensitivity) = 0.86/ (1 – 0.86) = **6.14**

Likelihood ratio negative = (1- sensitivity) / specificity = (1- 0.86) / 0 = **not calculable**

Positive post test probability = 99%

Negative post test probability = 0%

Copy of Consent email from James Serpell

From: James Serpell [mailto:serpell@vet.upenn.edu]
Sent: 07 March 2014 20:07
To: Hannah Plant
Subject: Re: C-BARQ Inquiry

Hi Hannah,

That sounds like a research project to me, so feel free to use the C-BARQ free of charge until you've finished your study. If Hearing Dogs decides to use the C-BARQ routinely in the future, they will be charged the same non-profit rate as all the other organizations that use it.

James

Copy of Consent email from Hearing Dogs For Deaf People

FW: MScRes Animal behaviour

JG

Jo Gray

Reply all

To:
dmills@lincoln.ac.uk;

Cc:
hzulch@lincoln.ac.uk;

Wed 27/03/2013 10:08

Dear Danny,

Following on from Hannah's email to you I wanted to confirm that all of the paperwork has been signed, and the Charity are fully supportive of Hannah's MSc. Financially, the Charity are contributing 50% of the cost of this degree and have made arrangements for Hannah to live on site whilst she completes the programme. We have also adjusted her working week to ensure that she has one day a week dedicated to her studies.

I look forward to seeing you in May where we can talk about the other collaborations that the future might hold for us. Exciting times!

Many thanks,
Jo